Communications Circuits: Analysis and Design

Clarke - Hess

* I *

[redacted]

PREFACE

This book has been written on the basis of a combined experience of more than thirty years of teaching about and working with electronic circuits of the type used in present-day communications and control systems. In this book we deal neither with semiconductor or vacuum tube manufacture nor with overall system design, but with the understanding and use of devices and configurations of devices that bridge the gap between these two disciplines. Although we do not deal particularly with the problems of integrated circuits, many of our results are indeed directly applicable to circuits in integrated form.

Chapter 1 offers a preview of things to come. Chapters 2 and 3 may be considered as a review of linear system concepts. Although the material stressed in these chapters ought to be presented in linear systems courses or textbooks, it has been our experience that the viewpoints that we find useful are often somewhat slighted there.

Chapters 4 and 5 provide the foundation for the rest of the book. Essentially they provide a reasonably rigorous but (we hope) intelligible account of both the small-signal and large-signal operation of both the single devices and the basic multiple device configurations that serve as the building blocks for all later circuits. These devices and configurations include the bipolar and field effect transistor, the differential pair, and the combination of resistance and reactance with these devices.

The approach taken allows one to make both large-signal and small-signal calculations without any ambiguity as to the resultant distortions or nonlinear by-products. While we did not invent all the results here, we have been using them and teaching them for some years. To our knowledge this is the first time that they have been coordinated and made available in one place.

Chapter 6 uses the vehicle of the sinusoidal oscillator to tie together all of the previous material. The techniques presented allow one to calculate the actual amplitude frequency and distortion of real oscillators rather than just to catalog a number of circuits. The squegging phenomenon in oscillators is treated in a unique and readily usable manner.

Chapter 7 considers the deliberate use of the device non-linearity to produce mixers and frequency converters. It explores the amplitude limitations upon "linear" mixing and the effect of deliberate or accidental series resistance upon the mixing process. This chapter also examines the feedthrough and the feedback problems involved in small signal RF amplifiers and AGC systems.

Chapter 8 is concerned with multipliers and amplitude modulators. It presents a step-by-step analysis of the popular Gilbert integrated four-quadrant multiplier as well as a number of other useful circuits. Chapter 9 discusses all types of power amplifiers from linear broadband Class A types through both tuned and broadband Class D types. Chapter 10 explores the amplitude demodulation problem in detail. It presents useful design results for the common narrowband peak envelope detector, which is usually used in circuits but rarely discussed in textbooks. Chapters 11 and 12 present a large amount of new material in their complete coverage of FM generation and detection.

Because the general principles of the first five chapters are applicable in some form to most of the circuits in the rest of the book, a unity is achieved that has often not been apparent in past books in this field. Thus instead of considering a seemingly endless variety of apparently different oscillators or detectors, one is able to group circuits into rather broad classes and show straightforward design or analysis procedures applicable to all of them.

Some of the early versions of this material were originally put into note form in 1962. All of it, except our last-minute revisions, has been used in various graduate and senior year courses at the Polytechnic Institute of Brooklyn. It is not reasonable to try to cover all this material in a one-semester course. Well-grounded students who can handle Chapters 2 and 3 by themselves, and who can absorb Chapters 4 and 5 in say three weeks, should be able to cover selected material from the remaining chapters without undue difficulty in a semester. A number of selections of coherent groups of material are possible. Most instructors should have no problem picking out a set that is both interesting to them and instructive to their students.

Homework problems are included at the end of each chapter. Illustrative examples are worked out in most chapters.

Our former colleagues and students at the Polytechnic Institute of Brooklyn deserve our thanks for their many stimulating criticisms and observations. Professors Gerald Weiss, Ronald Juels, and Marvin Panzer were particularly helpful in pointing out errors or areas in need of clarification. A special debt of gratitude is due to the various people who struggled with the typing and the drawings for the manuscript and the various sets of notes that preceded it.

As Department Head through much of the period that the book was in preparation, Professor Edward J. Smith and the rest of the administration of the Polytechnic were most kind in extending the use of various typing and reproduction facilities.

While book writing is never really a pleasure, it is exciting to find a simple way to solve a heretofore difficult problem. We have had many such exciting moments in preparing this book and we hope that the reader will be able to share some of our excitement as he uses it.

New York
May 1971

K. K. C.
D. T. H.

CONTENTS

Chapter 1 Preview

1.1 Basic circuit biasing 1
1.2 Wideband amplifier limits on "small-signal" operation 3
1.3 Narrowband amplifiers and limiters 6
1.4 Frequency multipliers 8
1.5 Mixers 8
1.6 Sine-wave oscillators 10
1.7 Conclusions 13

Chapter 2 Broadband and Narrowband Transformerlike Coupling Networks

2.1 Broadband transformer coupling 16
2.2 Parallel *RLC* circuit 25
2.3 Parallel *LC* circuit with series loss 32
2.4 Parallel resonant "transformerlike" networks 38
2.5 Parallel resonant transformers 48
2.6 Three-winding parallel resonant transformer 54
Appendix to Chapter 2: Transformer Equivalent Circuits 62

Chapter 3 Transmission of Signals Through Narrowband Filters

3.1 Low-pass equivalent networks for symmetrical bandpass networks . . 65
3.2 Impulse and step response 70
3.3 Narrowband networks with modulated inputs 72
3.4 Narrowband networks with periodic inputs 78
3.5 Total harmonic distortion 82
Appendix to Chapter 3: High-Q Filter Measurements 88

Chapter 4 Nonlinear Controlled Sources

4.1 General comments 90
4.2 Piecewise-linear source, single discontinuity 91
4.3 Multiple-segment piecewise-linear sources 94
4.4 Square-law characteristics 98
4.5 The exponential characteristic 104
4.6 The differential characteristic 114
4.7 Other gradual nonlinearities—pentodes 120
4.8 Effect of series resistance on the exponential characteristic 123
4.9 Clamp-biased square-law device 131
Appendix to Chapter 4: Fourier Expansions 144

Chapter 5 Reactive Element and Nonlinear Element Combinations

5.1 Capacitive coupling to nonlinear load . . . 149
5.2 Transient build-up to steady state . . . 159
5.3 Capacitively coupled transistor amplifier—constant current bias . . . 162
5.4 Capacitively coupled transistor amplifier—resistor bias . . . 169
5.5 Nonlinear loading of tuned circuits . . . 181
5.6 Transfer function for low-index AM input . . . 195

Chapter 6 Sinusoidal Oscillators

6.1 Operating frequency and minimum gain conditions for linear-feedback oscillators . . . 205
6.2 Amplitude-limiting mechanisms . . . 212
6.3 Frequency stability . . . 216
6.4 Self-limiting single-transistor oscillator . . . 222
6.5 Self-limiting differential-pair oscillator . . . 236
6.6 Self-limiting junction field effect transistor oscillators . . . 241
6.7 Crystal oscillators . . . 243
6.8 Squegging . . . 255
6.9 Bridge oscillators . . . 261
6.10 The one-port approach to oscillators . . . 268
6.11 The phase plane approach . . . 273
6.12 The distortion-operating frequency relationship . . . 279

Chapter 7 Mixers; RF and IF Amplifiers

7.1 The superheterodyne concept . . . 293
7.2 Mixer techniques . . . 295
7.3 Series resistance in mixers . . . 302
7.4 Practical mixer circuits . . . 305
7.5 Semiconductor converter circuits . . . 311
7.6 Tuned narrowband small-signal amplifiers . . . 314
7.7 Stages with double-tuned circuits . . . 328
7.8 Gain control circuits . . . 331
7.9 Noise, distortion, and cross modulation . . . 336
Appendix to Chapter 7: Comparisons of y-Parameters for Bipolar Transistors: Single-Ended, Differential-Pair, and Cascode . . . 345

Chapter 8 Amplitude Modulation

8.1 Amplitude modulation signals . . . 347
8.2 Amplitude modulation techniques . . . 353
8.3 Practical analog modulators and multipliers . . . 362
8.4 Practical chopper modulators . . . 376
8.5 Square-law modulator . . . 384
8.6 Tuned-circuit modulators . . . 387

Chapter 9 Power Amplifiers

9.1 "Ideal" power amplifiers—class A, single-ended . . . 401
9.2 Class B linear RF amplifiers . . . 405
9.3 Class C "linear" amplifiers . . . 408
9.4 RF class C amplifiers . . . 410
9.5 Narrowband class D power amplifiers . . . 415
9.6 Broadband class B amplifiers . . . 421
9.7 Broadband class D power amplifiers . . . 426
9.8 Practical power amplifiers . . . 432
9.9 High-level amplitude modulation . . . 447
Appendix to Chapter 9: Pulse Train Expansions . . . 454

De

Chapter 10 Amplitude Modulators

10.1 Amplitude demodulation techniques . . . 457
10.2 Practical average envelope detectors . . . 468
10.3 Narrowband peak envelope detector . . . 478
10.4 Practical narrowband peak envelope detector . . . 496
10.5 Broadband peak envelope detector . . . 498

Chapter 11 Generation of FM Signals

11.1 Frequency-modulated signals . . . 509
11.2 Transmission of FM signals through nonlinear networks . . . 515
11.3 Transmission of FM signals through linear filters . . . 521
11.4 Frequency modulation techniques—the FM differential equation . . . 526
11.5 Quasi-static frequency modulation . . . 532
11.6 Triangular-wave frequency modulation . . . 542
11.7 Practical square-wave frequency modulation . . . 553
11.8 Miscellaneous frequency modulators—the Armstrong method . . . 559
11.9 Frequency stabilization of frequency modulators . . . 562

De

Chapter 12 FM Modulators

12.1 Limiters . . . 571
12.2 Frequency-demodulation techniques . . . 577
12.3 Direct differentiation—the Clarke-Hess frequency demodulator . . . 586
12.4 Frequency-domain differentiation—the slope demodulator . . . 593
12.5 Time-delay differentiator, time-delay demodulator, Foster-Seeley demodulator, and ratio detector . . . 602
12.6 Pulse-count frequency demodulator . . . 618
12.7 More exotic FM detectors—the phase-locked loop, the frequency-locked loop, and the frequency demodulator with feedback . . . 623

Appendix: Modified Bessel Functions . . . 636

Answers to Selected Problems . . . 645

Index . . . 651

CHAPTER 1

PREVIEW

The purpose of this chapter is not to reduce the excitement of the book by "revealing the plot." We wish rather, by using one particular circuit as a vehicle, both to indicate a number of the techniques that we will explore later in detail and to demonstrate to the reader the power of these methods. We will show how we may rather easily get a clear insight into the design of such apparently diverse circuits as wideband small-signal amplifiers, large-signal narrowband amplifiers, frequency multipliers, active limiters, active mixers, and tuned-circuit sine-wave oscillators. By doing so, we hope to provide a framework for the general developments that follow and to share with others the enthusiasm that comes from being able to solve many heretofore difficult design and analysis problems.

In this chapter, because of its nature, we cannot develop all results or answer all questions. We trust that the unanswered questions will receive adequate treatment at a later point.

1.1 BASIC CIRCUIT BIASING

The circuit that we shall use as a skeleton upon which to construct our various examples is shown in Fig. 1.1–1.

This circuit is shown in the manner in which it might be constructed in integrated form. The sole purpose of the lower two transistors is to provide a constant current

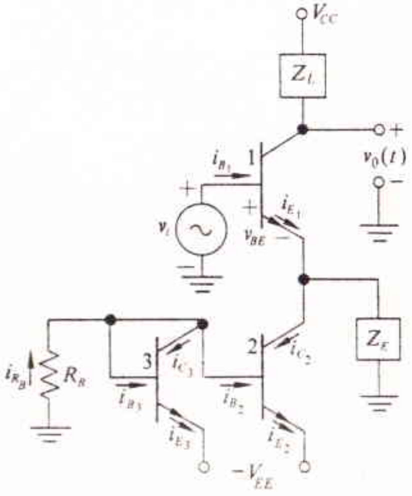


Fig. 1.1-1 Basic junction transistor amplifier.

bias source for transistor 1. (Transistor 3 might be viewed as a diode; however, in integrated circuits diodes are normally constructed as transistors.)

Our key assumption is that the emitter current and the base-emitter voltage of the transistors are related by Eq. (1.1–1)†:

$$i_E = I_{ES}e^{v_{BE}q/kT}, \tag{1.1–1a}$$

$$v_{BE} = \frac{kT}{q}\ln\frac{i_E}{I_{ES}}, \tag{1.1–1b}$$

where $k = 1.38 \times 10^{-23} J/°\text{K}$ is Boltzmann's constant, $q = 1.6 \times 10^{-19}$ C is the electronic charge, and I_{ES} is the emitter saturation current.

Let us make a further set of assumptions: that $i_C = \alpha i_E$ and $i_B = (1 - \alpha)i_E$, and that α is both close to unity and independent of i_E. (The assumption of a constant alpha is rarely true if i_E varies over a wide range; however, if alpha approaches unity, then this variation is normally a second-order effect.)

Since I_{ES} is of the order of 2×10^{-16} A for small silicon integrated circuit transistors and since $kT/q \approx 26$ mV at normal room temperatures ($T = 300°\text{K}$), Eq. (1.1–1b) may be employed to determine the required values of v_{BE} (or V_{BE} for the case of a bias voltage) to produce various values of i_E (or I_E). Several values of V_{BE} vs. I_E are presented in Table 1.1–1. It is apparent that V_{BE} varies only slightly for large variations in I_E; hence in many applications V_{BE} may be approximated by a constant of approximately $\frac{3}{4}$ V.

Table 1.1–1 Value of V_{BE} required for various values of I_E

V_{BE}, mV	I_E, mA
700	0.1
760	1
820	10
880	100

As connected in Fig. 1.1–1, transistors 2 and 3 must have the same value for v_{BE} (or V_{BE}). If they occupy the same area and are on the same chip, they will have almost identical values for I_{ES}. Therefore, $i_{E2} = i_{E3}$ or, for biasing purposes, $I_{E2} = I_{E3}$. Now $I_{R_B} = (V_{EE} - V_{BE})/R_B$. If V_{BE} is approximated by 3/4 V (so long as $V_{EE} \gg V_{BE}$ this is reasonable), then I_{R_B} is known. However, $I_{R_B} = I_{E3} + (1 - \alpha)I_{E2}$ or

$$I_{E2} = \frac{V_{EE} - 0.75}{(2 - \alpha)R_B} \approx \frac{V_{EE} - 0.75}{R_B}, \tag{1.1–2}$$

---

† A somewhat more accurate representation would be

$$i_E = I_{ES}\,e^{q\gamma v_{BE}/kT},$$

where $\frac{1}{2} \leq \gamma \leq 1$ depending on the transistor material, i.e., germanium or silicon. In any situation which warrants it γ may be included without affecting any of the derived results.

and thus

$$I_{C2} = \alpha I_{E2} \approx \frac{V_{EE} - 0.75}{R_B}. \tag{1.1–3}$$

So long as Z_E contains a series capacitor (no dc path), then $I_{E1} = I_{C2}$ and the upper transistor is biased at a constant current level.

1.2 WIDEBAND AMPLIFIER LIMITS ON "SMALL-SIGNAL" OPERATION

Let us first consider the case where Z_L is a resistor R_L, Z_E is a capacitor C_E, $v_i = V_1 \cos \omega t$, $1/\omega C_E$ approaches an ac short circuit, and ω is low enough so that transistor reactances may be ignored. We assume that V_{CC} and V_{EE} are large enough so that the collector-base junctions of both transistors 1 and 2 always stay reverse biased.

Since C_E is an ac short circuit, v_i appears directly across the emitter-base junction of transistor 1. In addition, any dc voltage V_{dc} which is developed across C_E appears across the junction; hence $v_{BE1} = v_i + V_{\text{dc}}$. When v_i is zero, i_E is forced to be equal to I_{C2}; hence

$$V_{\text{dc}\,Q} = \frac{kT}{q}\ln\frac{I_{C2}}{I_{ES}}.$$

(The subscript Q denotes the quiescent value of a parameter.)

For the case where v_i is not equal to zero, Eq. (1.1–1a) may be employed to obtain

$$\begin{aligned} i_E &= I_{ES}[e^{V_{\text{dc}}q/kT}]e^{(V_1 q/kT)\cos \omega t} \\ &= I_{ES}[e^{V_{\text{dc}}q/kT}]e^{x \cos \omega t}, \end{aligned} \tag{1.2–1}$$

where $x = V_1 q/kT$ to normalize the drive voltage. Now from a known Fourier series expansion,

$$e^{x \cos \omega t} = I_0(x) + 2\sum_{1}^{\infty} I_n(x) \cos n\omega t, \tag{1.2–2}$$

where $I_n(x)$ is a modified Bessel function of the first kind, of order n and argument x. (Properties of these tabulated functions as well as further references concerning them will be found in the Appendix at the back of the book.) The modified Bessel functions are all monotonic and positive for $x \geq 0$ and $n \geq 0$; $I_0(0)$ is unity, whereas all higher-order functions start at zero. As $x \to 0$,

$$I_n(x) \to \frac{(x/2)^n}{n!},$$

when n is a positive integer.

Combining Eqs. (1.2–1) and (1.2–2), we obtain

$$i_E = I_{ES}e^{V_{\text{dc}}q/kT} I_0(x)\left[1 + 2\sum_{1}^{\infty} \frac{I_n(x)}{I_0(x)} \cos n\omega t\right]. \tag{1.2–3}$$

It is apparent from Eq. (1.2–3) that the average (or dc) value of i_E is given by

$$\bar{i}_E = I_{ES} e^{V_{dcq}/kT} I_0(x). \tag{1.2–4}$$

However, the biasing circuitry demands that $\bar{i}_E = I_{C2}$; hence i_E may be written in the simplified form

$$i_E = I_{C2}\left[1 + 2\sum_{1}^{\infty} \frac{I_n(x)}{I_0(x)} \cos n\omega t\right]. \tag{1.2–5}$$

In addition, V_{dc} may be obtained from Eq. (1.2–4) to be of the form

$$V_{dc} = \frac{kT}{q} \ln \frac{I_{C2}}{I_{ES} I_0(x)} = \frac{kT}{q} \ln \frac{I_{C2}}{I_{ES}} - \frac{kT}{q} \ln I_0(x)$$

$$= V_{dcQ} - \frac{kT}{q} \ln I_0(x). \tag{1.2–6}$$

Table 1.2–1 presents several sets of data concerning the modified Bessel functions that will be of interest to us. From the first column of this table we see that if $V_1 =$ 260 mV, so that $x = 10$, then the dc voltage shifts by 206 mV from its Q-point value. We can also see from the other columns that the peak value of the fundamental component of the collector current of transistor 1 is $1.9 I_{C2}$, while the percentage second-harmonic distortion in this current is 85%.

Table 1.2–1

x	$\ln I_0(x)$	$\frac{2I_1(x)}{I_0(x)}$	$\frac{I_2(x)}{I_1(x)}$
0	0.000	0.000	0.000
0.5	0.062	0.485	0.124
1	0.236	0.893	0.240
2	0.823	1.396	0.433
5	3.30	1.787	0.719
10	7.93	1.897	0.854
20	17.6	1.949	0.926

Apparently a 260 mV peak sinusoidal signal is not a small signal at all from the viewpoint of this amplifier. The limits of small-signal operation are made clearer by a study of Figs. 1.2–1 and 1.2–2. Figure 1.2–1 shows that the output fundamental is only roughly linearly proportional to the input voltage, or equivalently x, for $x \leq 1$. However, to keep $I_2(x)/I_1(x)$, which is the percent second-harmonic distortion, below .025 ($2\frac{1}{2}$% distortion), it is necessary to keep x below 0.1.† Consequently, for small-signal operation $V_1 \leq 2.6$ mV or equivalently $|v_1| \leq 2.6$ mV.

It is apparent from Eq. (1.2–1) that the emitter current and, in turn, the collector current of transistor 1 are proportional to $e^{x \cos \omega t}/e^x$ for any fixed value of x. (We incorporate the e^x term in the denominator for normalization purposes only.) Consequently, the plot of $e^{x \cos \omega t}/e^x$ shown in Fig. 1.2–2 yields a normalized picture of the collector current as a function of time over one cycle of the input voltage $v_1 = V_1 \cos \omega t$. Clearly by the time $x = 10$, the collector current is flowing in narrow pulses approximately $\frac{1}{4}$ cycle wide; hence the operation of the amplifier is certainly not linear. In fact, as x increases above one, the overall current waveshape rapidly ceases to be cosinusoidal. For larger values of x the dc bias shift effectively aids the signal in holding the base-emitter junction off for a good portion of the cycle.

---

† For small values of x, $I_2(x)/I_1(x) \approx x/4$ [cf. Eq. (A–2) in the Appendix at the back of the book].

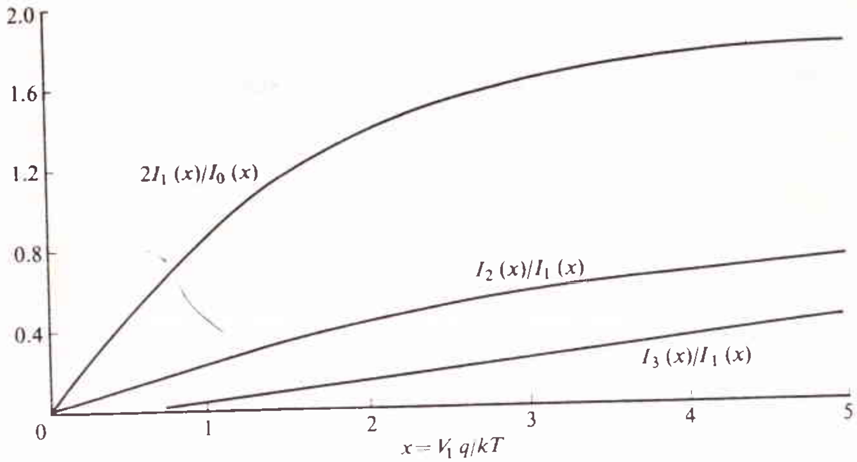


Fig. 1.2–1 Functions of modified Bessel functions vs. the normalized parameter x.

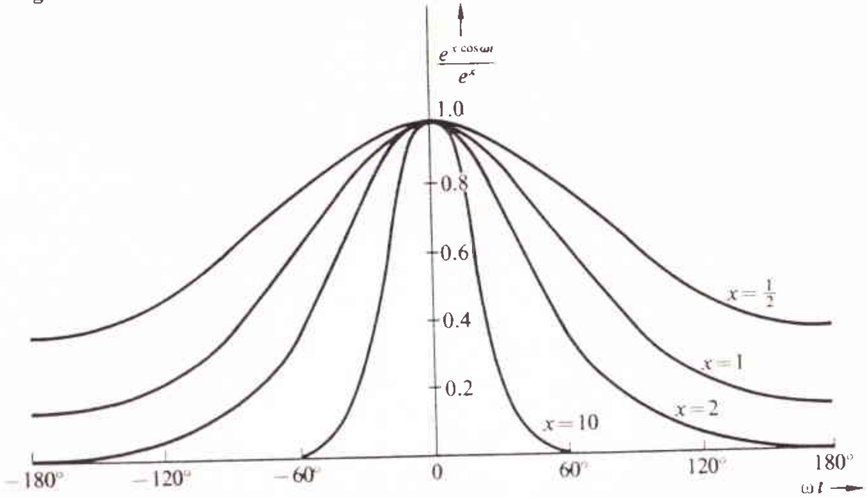


Fig. 1.2–2 Normalized collector currents vs. angle for exponential junction driven by sine wave.

With $|v_i| \leq 2.6$ mV the output voltage of the amplifier takes the form

$$v_o(t) = V_{CC} - i_C R_L = V_{CC} - \alpha I_{C2} R_L - \alpha I_{C2} R_L \frac{2I_1(x)}{I_0(x)} \cos \omega t. \tag{1.2–7}$$

For small values of x, however, $2I_1(x)/I_0(x) \approx x = V_1q/kT$; consequently,

$$v_o(t) = V_{CC} - \alpha I_{C2}R_L - g_{mQ}R_LV_1 \cos \omega t, \tag{1.2–8}$$

where $g_{mQ} = \alpha I_{C2}q/kT$ is defined as the transconductance of the transistor. Note that the value of g_m is exactly that value which would be obtained as the incremental ratio of collector current to base-emitter voltage evaluated about the Q-point; that is,

$$\left.\frac{\partial i_C}{\partial v_{BE}}\right|_{i_C=\alpha I_{C2}} = \alpha \left.\frac{\partial i_E}{\partial v_{BE}}\right|_{i_E=I_{C2}} = \frac{\alpha q I_{C2}}{kT} = g_{mQ}, \tag{1.2–9}$$

where $i_E = I_{ES}e^{v_{BE}q/kT}$. Thus for $|v_i| < 2.6$ mV, classical small-signal analyses may be employed.

As we shall see in a later chapter, one way to extend the broadband linear signal handling capacity of a transistor amplifier is to include an unbypassed emitter resistor. In the circuit under discussion this resistor R_E would be placed in series with C_E. Such a resistor reduces the fundamental gain of the stage by a factor of

$$\frac{1}{1 + R_E(I_{C2}q/kT)} = \frac{1}{1 + g_{mQ}R_E},$$

where g_{mQ} is the small-signal Q-point transconductance with R_E shorted.

The effect of this series resistance is to linearize the characteristic so that, while it does reduce the fundamental gain, it reduces the harmonic distortion even more rapidly.

As a practical matter one should note that all of the foregoing discussion would be unchanged if the v_i generator were included in series with C_E while the base of transistor 1 was grounded.

1.3 NARROWBAND AMPLIFIERS AND LIMITERS

A different approach to utilizing the circuit of Fig. 1.1–1 would be to again let Z_E be a single capacitor C_E and let Z_L be a parallel RLC circuit tuned to the frequency ω of the input signal as shown in Fig. 1.3–1. For a parallel RLC circuit, the magnitude of the impedance at the fundamental frequency, $Z_L(j\omega)$, is, in general, greater than the magnitude of the impedance at the nth harmonic, $|Z_L(jn\omega)|$; in particular,

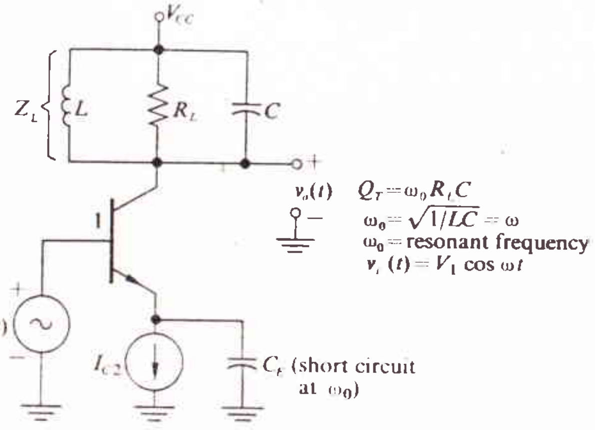

V_{CC}
Z_L
L
R_L
C
+
$v_o(t)$
−
1
$v_i(t)$
+
−
I_{C2}
C_E (short circuit at ω_0)
$Q_T = \omega_0 R_L C$
$\omega_0 = \sqrt{1/LC} = \omega$
ω_0 = resonant frequency
$v_i(t) = V_1 \cos \omega t$

Fig. 1.3–1 Circuit of Fig. 1.1–1 with Z_L replaced by a tuned circuit.

$$\frac{|Z_L(jn\omega)|}{|Z_L(j\omega)|} \approx \frac{n}{(n^2 - 1)Q_T}, \tag{1.3–1}$$

where Q_T is the Q of the resonant circuit. Therefore, if Q_T is sufficiently high, we can obtain almost a pure sine-wave output voltage $v_o(t)$ in spite of large harmonic components in the collector current of transistor 1.

For example, if $x = 5$, then $Q_T = 48$ reduces the second-harmonic output voltage component to 1% of the fundamental and the third-harmonic voltage component to 0.31% (the collector current has distortion components of 72% and 40% respectively). Therefore, as a good approximation, the output voltage $v_o(t)$ may be written as

$$v_o(t) = V_{CC} - \alpha I_{C2}R_L\frac{2I_1(x)}{I_0(x)} \cos \omega t, \tag{1.3–2}$$

where $R_L = Z_L(j\omega)$ is the impedance of the parallel RLC circuit at resonance. Since the impedance of the resonant circuit to dc is zero, no dc voltage is built up across Z_L.

In this case, instead of the small-signal transconductance g_{mQ}, it is convenient to define a large-signal average transconductance G_m which is equal to the ratio of the fundamental collector current I_{C11} to the fundamental driving voltage V_1:

$$G_m = G_m(x) = \frac{I_{C11}}{V_1} = \frac{\alpha I_{C2}}{V_1}\frac{2I_1(x)}{I_0(x)} = g_{mQ}\frac{2I_1(x)}{xI_0(x)}. \tag{1.3–3}$$

$x = \frac{V_1}{kT}$

With this definition for $G_m(x)$, $v_o(t)$ may be written in the equivalent form

$$v_o(t) = V_{CC} - G_m(x)R_LV_1 \cos \omega t, \tag{1.3–4}$$

which is similar in form to the output of the small-signal amplifier. The basic difference is that $G_m(x)$ is a function of V_1 (or x) and no longer a constant.

Figure 4.5–6 presents values for $G_m(x)/g_{mQ}$ for various values of x. From Fig. 4.5–6 we see that G_m is down 1 dB from its $x = 0$ value when $x = 1$; hence, though the harmonic distortion has been removed, the amplifier can operate only in an approximately "linear" fashion with input amplitudes below 26 mV peak. By linear, in this case, we mean that there is a constant ratio between input and output signal levels, and that this ratio is independent of signal level; this is necessary if an AM wave is to be amplified. If we wish to handle larger input signals in a linear manner, then the unbypassed emitter resistor again provides the means.

If, on the other hand, we want to remove amplitude variations in V_1 from the output, i.e., if we wish to produce a "limiter," then we need only increase x. From Table 1.2–1 or from Fig. 1.2–1 we note that as x increases, $2I_1(x)/I_0(x)$ approaches a saturation value of 2; hence $v_o(t)$ given by Eq. (1.3–2) reduces to

$$v_o(t) = V_{CC} - \alpha 2I_{C2}R_L \cos \omega t, \tag{1.3–5}$$

which is clearly independent of variations in V_1.

As an example, we consider the case where V_1 varies between 130 mV and 520 mV (x varies between 5 and 20) because of a spurious amplitude modulation. If we define the modulation index as

$$m = \frac{V_{1_{\max}} - V_{1_{\min}}}{V_{1_{\max}} + V_{1_{\min}}},$$

then the input modulation index is $m = 0.6$ (or 60%). Since for $x = 5$, $2I_1(x)/I_0(x) = 1.787$, and for $x = 20$, $2I_1(x)/I_0(x) = 1.949$, and since the amplitude of the ac component of $v_o(t)$ is proportional to $2I_1(x)/I_0(x)$ [cf. Eq. (1.3–2)], the output modulation index is

$$m_o = \frac{1.949 - 1.787}{1.949 + 1.787} = 0.0435 \quad \text{(or 4.35\%)}.$$

A further stage driven with this signal at a normalized level such that $x \geq 10$ could reduce the output modulation below 0.05%.

1.4 FREQUENCY MULTIPLIERS

As we saw in Figs. 1.2–1 and 1.2–2, as x increases, the harmonic component of the collector current increases. For $x = 10$, $I_2(x)/I_1(x) = 0.85$, $I_3(x)/I_1(x) = 0.66$, $I_4(x)/I_1(x) = 0.46$, and $I_5(x)/I_1(x) = 0.29$. Therefore if we tune the output-tuned circuit to a harmonic of the input, we can obtain an appreciable voltage at least up to the fifth harmonic for an input drive of 260 mV ($x = 10$). (For $x = 20$, $I_5(x)/I_1(x)$ has increased to 0.54.) Such circuits are known as frequency multipliers. They are widely used to obtain a higher frequency from a stable crystal oscillator or, in FM systems, to increase the output FM deviation. Specifically, if the parallel RLC circuit is tuned to the nth harmonic of the input, $v_o(t)$ is given by

$$v_o(t) = V_{CC} - \alpha I_{C2} R_L \frac{2I_n(x)}{I_0(x)} \cos n\omega t. \tag{1.4–1}$$

1.5 MIXERS

So far we have driven the junction of transistor 1 with a single-frequency cosinusoid. Let us now consider the case where $v_i(t) = V_1 \cos \omega_1 t + g(t) \cos \omega_2 t$. The signal at frequency ω_1 may be thought of as a local oscillator signal in a superheterodyne receiver; $g(t) \cos \omega_2 t$ may be thought of as a low-level received amplitude-modulated (AM) signal which we wish to translate to the intermediate frequency (IF) of the receiver. If we again note that for transistor 1 in Fig. 1.1–1 (with $Z_E = C_E$) the base-emitter voltage is given by $v_{BE} = v_i + V_{\mathrm{dc}}$, then we may write the emitter current in the form

$$i_E = I_{ES} e^{V_{\mathrm{dc}} q/kT} e^{x \cos \omega_1 t} e^{[qg(t)/kT] \cos \omega_2 t}. \tag{1.5–1}$$

If we assume $|g(t)| \leq 2.6$ mV, then $e^{[qg(t)/kT] \cos \omega_2 t}$ may be approximated by $1 + [qg(t)/kT] \cos \omega_2 t$. In addition, if we replace $e^{x \cos \omega_1 t}$ by its Fourier series, Eq. (1.5–1) simplifies to

$$i_E = I_{ES} e^{V_{\mathrm{dc}} q/kT} I_0(x) \left[1 + \frac{2I_1(x)}{I_0(x)} \cos \omega_1 t + \frac{2I_2(x)}{I_0(x)} \cos 2\omega_1 t + \cdots \right]$$
$$\times \left[1 + \frac{qg(t)}{kT} \cos \omega_2 t \right]. \tag{1.5–2}$$

Finally, by noting that $\cos A \cos B = \frac{1}{2}[\cos (A - B) + \cos (A + B)]$, we may rewrite i_E in the form

$$i_E = I_{C2} \left[1 + \frac{2I_1(x)}{I_0(x)} \cos \omega_1 t + \frac{2I_2(x)}{I_0(x)} \cos \omega_2 t + \cdots \right] + g(t) \frac{qI_{C2}}{kT}$$
$$\times \left[\cos \omega_2 t + \frac{I_1(x)}{I_0(x)} \cos (\omega_1 - \omega_2)t + \frac{I_1(x)}{I_0(x)} \cos (\omega_1 + \omega_2)t + \cdots \right]. \tag{1.5–3}$$

Hence we have generated AM waves with envelopes proportional to the input envelope $g(t)$ at frequencies $\omega_1 - \omega_2$, $\omega_1 + \omega_2$, $2\omega_1 - \omega_2$, $2\omega_1 + \omega_2$, etc.

If we now choose Z_L as a parallel RLC circuit tuned at $\omega_1 - \omega_2$ with a value of Q_T sufficient to remove other frequency components from the output [but not so large that the envelope information of $g(t)$ is filtered], then the output takes the form

$$v_o(t) = V_{CC} - R_L \frac{\alpha q I_{C2}}{kT} \frac{I_1(x)}{I_0(x)} g(t) \cos (\omega_1 - \omega_2)t$$
$$= V_{CC} - g_c R_L g(t) \cos (\omega_1 - \omega_2)t,$$

where

$$g_c = \frac{\alpha q I_{C2}}{kT} \frac{I_1(x)}{I_0(x)} = g_{mQ} \frac{I_1(x)}{I_0(x)}.$$

Clearly the input AM wave has been translated in frequency from ω_2 to $\omega_1 - \omega_2$. By choosing the oscillator frequency ω_1 correctly we can shift (or mix) the input AM wave to any desired intermediate frequency.

The quantity g_c, which may be interpreted as the ratio of the envelope of the collector current at the frequency $\omega_1 - \omega_2$ to the envelope of the input voltage at frequency ω_2, is called the *conversion transconductance*. Since $I_1(x)/I_0(x)$ increases monotonically toward an asymptote of unity for large values of x (or equivalently V_1), it is apparent that g_c is optimized by choosing V_1 greater than 260 mV ($x > 10$). For this case $g_c \approx g_{mQ}$ and the mixer not only translates in frequency but also amplifies.

As an example of this fact we consider the case in which $g(t) = (1\text{ mV})(1 + \cos \omega_m t) \cos \omega_2 t$ (where $\omega_m \ll \omega_2$), $R_L = 10\text{ k}\Omega$, $I_{C2} = 2.6$ mA, $V_1 = 260$ mV, $\alpha \approx 1$, and Z_L is a parallel RLC circuit tuned at $\omega_1 - \omega_2$. Clearly then for this circuit

$g_c = (0.1)(0.948)$ mho $= 0.0948$ mho. Consequently, the output voltage is

$$v_o(t) = V_{CC} - (0.948\ \text{V})(1 + \cos\omega_m t)\cos(\omega_1 - \omega_2)t.$$

The output signal is shifted in frequency and amplified by a factor of almost 1000.

It is interesting that mixers of this form are employed in all superheterodyne receivers, that is, in more than 99% of the world's receivers of any kind.

1.6 SINE-WAVE OSCILLATORS

To operate the mixer we required a local oscillator; hence every superheterodyne receiver requires an oscillator. At the same time every transmitter also requires an oscillator. The first-order characteristics of an oscillator are its waveshape, its frequency, and its amplitude. Second-order characteristics are the frequency and amplitude stability with changes in time, temperature, voltage, and physical movement.

To set the frequency of a sine-wave oscillator we connect it into a feedback loop so that positive feedback of exactly 360° is possible only at the desired frequency. To build frequency stability into it we concentrate most of the phase shift vs. frequency dependence into one portion of the circuit (often a quartz crystal or a high-Q tuned circuit). The oscillator often attains its desired amplitude by reaching a balance between the feedback allowed by the passive portions of the circuit and the nonlinear gain offered by the active portion of the circuit (the transistor in the case we are about to consider).

Figure 1.6–1 shows a sine-wave oscillator circuit constructed from the basic circuit of Fig. 1.1–1. For this circuit a 360° phase shift around the loop is possible only in the vicinity of the tuned circuit resonant frequency ω_0; hence, if an oscillation occurs, it has a frequency of approximately ω_0. Let us now assume that the resonant circuit has a high Q_T; then, if the circuit oscillates, the voltage across it is almost sinusoidal even if the collector current flows in narrow pulses. If we assume also that this voltage, $v_t = V_t \cos\omega_0 t$, is stepped down by the capacitance ratio $n = C_1/(C_1 + C_2)$ (cf. Chapter 2), then a sinusoidal drive voltage of the form $V_1 \cos\omega_0 t$ appears at the emitter of transistor 1, where $V_1 = nV_t$. This emitter voltage, in general, causes a nonlinear, pulselike collector current.

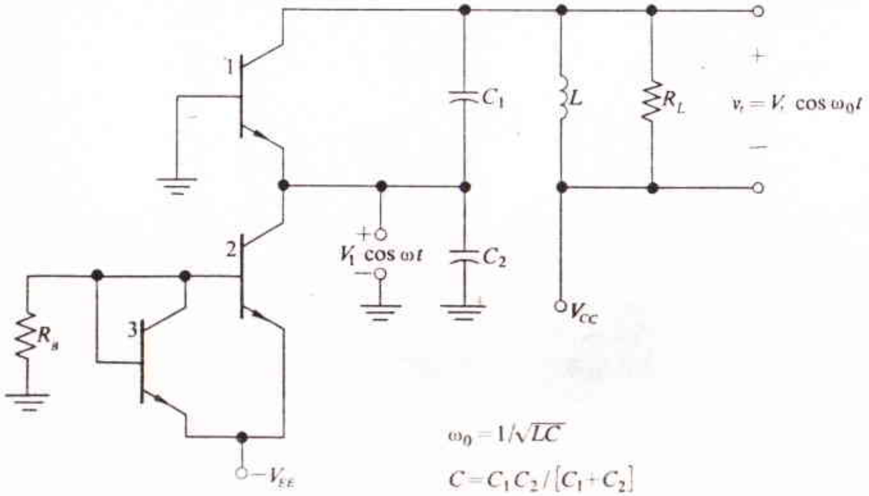


Fig. 1.6–1 Sine-wave oscillator.

We demonstrate in Section 5.5 that the loading of the transistor emitter junction upon C_2 is equivalent to a resistance of $\alpha/G_m(x)$. In addition, we demonstrate that this loading may be reflected across the inductor L as a conductance of $n^2[G_m(x)/\alpha]$, where n is again $C_1/(C_1 + C_2)$. Consequently, the total effective conductance appearing across the inductor is

$$G_T = G_L + \frac{n^2 G_m(x)}{\alpha}.$$

According to the Barkhausen criterion, for a sustained sinusoidal oscillation at ω_0, $A_L(j\omega_0) = 1$, where $A_L(j\omega)$ is the loop gain. To evaluate the loop gain we break the loop at the emitter, apply a signal of the form $V_1 \cos\omega_0 t$ to the emitter, terminate the broken loop in a resistance of $\alpha/G_m(x)$, and determine the signal across the termination of the loop. The broken loop is shown in Fig. 1.6–2. The capacitor C_E, which has no effect on the calculation of $A_L(j\omega_0)$ since it is an ac short circuit, is incorporated to preserve the dc bias conditions.

Now, with the loop broken, the oscillator reduces to a narrowband amplifier for which we may write

$$v_t = \frac{V_1 G_m(x)\cos\omega_0 t}{G_L + \dfrac{n^2 G_m(x)}{\alpha}}. \tag{1.6–1}$$

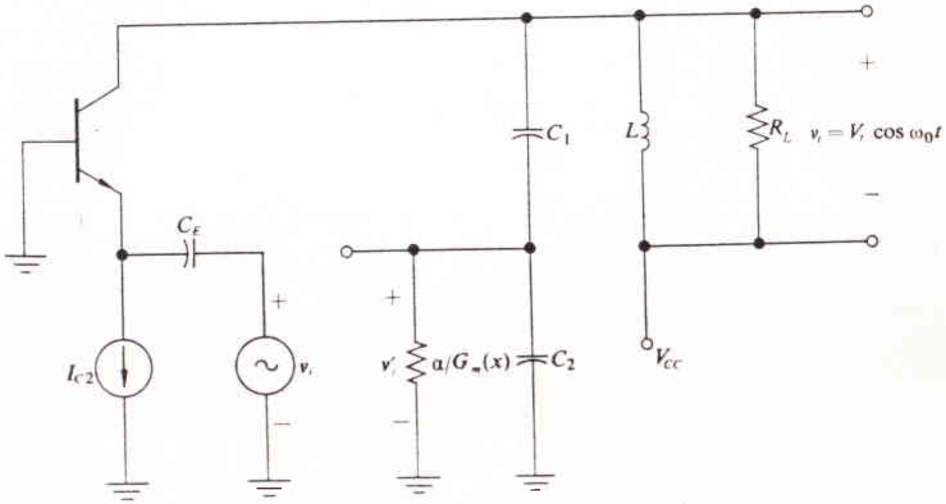


Fig. 1.6–2 Circuit for determining loop gain.

Since the two capacitors act as a step-down transformer with ratio n,

$$v_i' = \frac{nV_1 G_m(x) \cos \omega_0 t}{G_L + \dfrac{n^2 G_m(x)}{\alpha}} \tag{1.6–2}$$

and

$$A_L(j\omega_0) = \frac{v_i'}{v_i} = \frac{nG_m(x)}{G_L + \dfrac{n^2 G_m(x)}{\alpha}}. \tag{1.6–3}$$

Stable sinusoidal oscillations occur at ω_0 for

$$\frac{nG_m(x)}{G_L + \dfrac{n^2 G_m(x)}{\alpha}} = 1$$

or equivalently

$$G_m(x) = \frac{G_L}{n(1 - n/\alpha)}. \tag{1.6–4}$$

Equation (1.6–4) specifies the value of $G_m(x)$ required by the passive portion of the circuit. The amplitude must now adjust itself so that the transistor supplies this $G_m(x)$. If I_{C2} is known, then $g_{mQ} = I_{C2}q/kT$ follows, and from Fig. 4.5–6 we may determine the x that corresponds to the required G_m.

For example, if $C_1 = 100$ pF, $C_2 = 11{,}200$ pF, and $R_L = 13.7$ kΩ, then

$$n = 0.00885 \qquad \text{and} \qquad G_m(x) \approx \frac{G_L}{n} = 8300\ \mu\text{mho}.$$

If, in addition, $I_{C2} = 0.5$ mA, then

$$g_{mQ} = 19{,}200\ \mu\text{mho} \qquad \text{and} \qquad G_m(x)/g_{mQ} = 0.432;$$

hence from Fig. 4.5–6,

$$x \approx 4, \qquad V_1 = 4 \times 26 = 104\ \text{mV}, \qquad \text{and} \qquad V_t = \frac{V_1}{n} = \frac{104\ \text{mV}}{0.00885} = 11.8\ \text{V}.$$

It is quite obvious that V_{CC} must exceed 12 V if collector-base saturation is not to occur in transistor 1. So long as $V_{CC} > 12$ V, the previous amplitude is the amplitude at which the circuit stabilizes. In addition, in this case,

$$Q_T = \omega_0 \frac{C_1 C_2}{C_1 + C_2} R_T \approx \omega_0 C_1 R_L \approx 13.7,$$

which is not nearly as high as we would normally want the Q of an oscillator to be. However, even in this relatively low-Q case we have only 5% second-harmonic voltage across the tuned circuit; hence our assumption of a pure sine-wave drive

---

was not bad. We shall show later that such a circuit oscillates within two parts in a thousand of the nominal center frequency of the tuned circuit alone.

1.7 CONCLUSIONS

Now that we have seen some of the possibilities of this simple circuit, we shall go back and examine both a number of passive circuits and a number of other nonlinearities in detail. Then we shall return to explore in more depth each of the circuits discussed here, and also to discover many other circuits of immediate interest to the communication or control system designer. Before we plunge into nonlinear controlled sources and then into circuits, we devote a chapter to passive transformerlike networks and a chapter to the response of narrowband filters to modulated signals. There are no review chapters on basic electronics, for instance, on biasing small-signal amplifiers. For readers who feel deficient in such areas, some suggested background reading is listed below.

SUGGESTED BACKGROUND READING IN ELECTRONICS

Angelo, E. J., Jr., *Electronics: FET's, BJT's and Microcircuits*, McGraw-Hill, New York (1969).

Gray, P. E., *Introduction to Electronics*, John Wiley, New York (1967). A 325-page paperback developed for an introductory course in electronics. About half of the book is devoted to physics, diodes, and diode circuits, one-third to junction transistors, and one-sixth to field effect transistors and vacuum tubes.

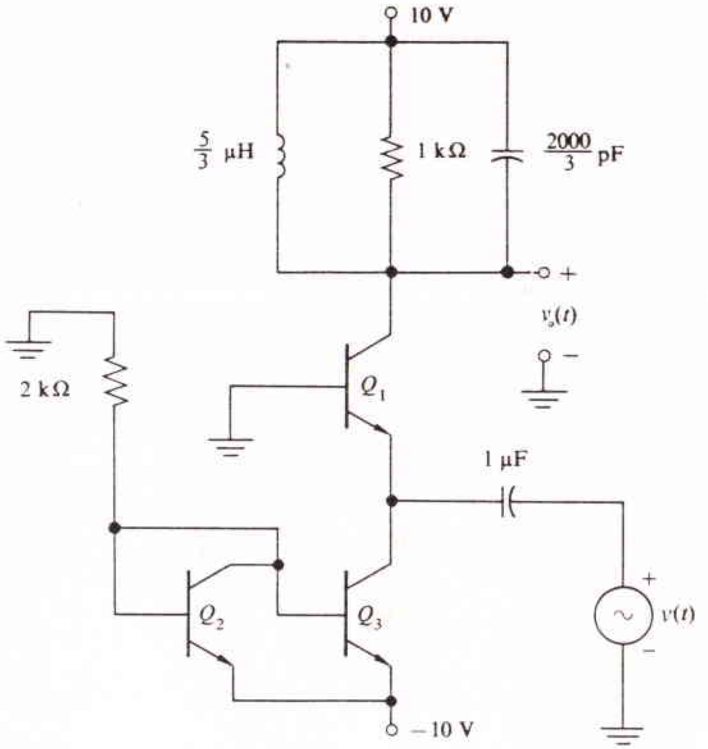
PROBLEMS

Problems 1.1 through 1.5 are all based on the circuit of Fig. 1.1–1.

1.1 Supposing that $|V_{EE}| = 3$ V, $R_B = 3$ kΩ, all alphas $= 0.98$ and all transistors are identical, silicon, and have $I_{ES} = 2 \times 10^{-16}$ A, find I_{C1}. If $V_{CC} = +10$ V, then determine the value of R_L that can be used to replace Z_L so that the output dc voltage level will be $+5$ V. What is the approximate power dissipation in each transistor for this case?

1.2 Suppose Z_E is replaced by an ac short circuit in the circuit of Problem 1.1. Sketch v_o for the cases where v_i is a pure sine wave having peak amplitudes of 1 mV, 2.6 mV, 26 mV, and 260 mV. (In Section 5.3 this case is considered in detail: only a reasonable estimate of the output is required at this point.)

1.3 Suppose Z_E is replaced by a 100 Ω resistor in series with an ac short circuit. Repeat Problem 1.2 for the cases in which v_i has a peak amplitude of 1 mV and of 260 mV. Compare the



$v(t) = (260\text{ mV})\cos 10^7 t$
$T = 27°\text{C}, \quad \beta = 98$
$V_{BE} = 0.7\text{ V}$

Figure 1.P–1

results with the previous problem. (The second case is not trivial; the solution is covered in detail in Chapter 5.)

1.4 Repeat Problem 1.2 for the case where R_L is shunted by a parallel LC combination tuned to the resonant frequency of the input sinusoidal signal. Does saturation occur?

1.5 Repeat Problem 1.2 for the case where R_L is shunted by a parallel LC combination tuned to the second harmonic of the input sinusoidal signal and compare the results with those of Problem 1.4.

1.6 For the circuit shown in Fig. 1.P–1, determine an expression for $v_o(t)$ (Q_2 and Q_3 are identical).

1.7 For the circuit shown in Fig. 1.P–2, determine the quiescent values of i_{E1}, v_{EB1}, and v_o when $I_{ES1} = 10^{-13}$ A, $I_{ES2} = 2 \times 10^{-13}$ A, and $I_{ES3} = 1.5 \times 10^{-13}$ A.

1.8 For the circuit of Fig. 1.P–2, determine $v_o(t)$ where $v_i = (1\text{ mV})\cos 10^6 t$ and $I_{ES1} = I_{ES2} = I_{ES3} = 10^{-13}$ A.

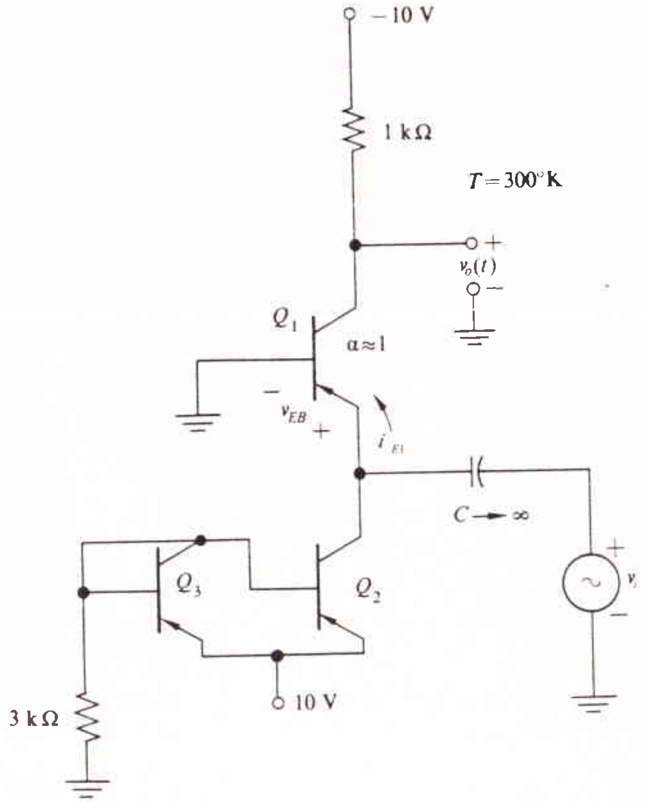


Figure 1.P–2

CHAPTER 2

BROADBAND AND NARROWBAND TRANSFORMERLIKE COUPLING NETWORKS

In this chapter we explore the similarities among a number of passive networks all of which have widespread practical application. All of these networks have the property of being able to transform impedance levels and hence voltage and current levels. Initially we consider a broadband transformer, and in later sections we show how a number of practical circuits may be reduced to the combination of a parallel *RLC* circuit and an ideal transformer. Throughout the chapter, emphasis is placed on plausible approximations, usually based on a consideration of the pole-zero diagram for the circuit in question.

The reader may question the necessity of such a chapter, since he has undoubtedly already had one or more courses in network theory or linear circuits. We have included the chapter because it has been our experience that such courses—or the textbooks used in them—rarely bring out the similarities in the circuits discussed here or make evident the approximations that suffice to simplify them. It is our aim in later chapters to combine these circuits with various nonlinear elements to make useful circuits. Before undertaking this combination it seems wise to have a thorough familiarity with the individual pieces.

The reader eager to get on to complete circuits might examine the equivalences shown in Table 2.5–1 and the illustrative examples at the end of Section 2.4. If these are all "old hat," then we urge him to push on; if not, we recommend this chapter as a foundation for later work.

2.1 BROADBAND TRANSFORMER COUPLING

In this section we study the frequency and time-domain properties of a linear network consisting of a resistive load coupled to a driving voltage source by means of a broadband transformer as shown in Fig. 2.1–1. Such networks are useful for providing dc isolation and the possibility of phase inversion between the input and the output; they are also employed when the load resistor must be scaled in value to "match" the driver over a broad band of frequencies. For example, a transistor power amplifier might require a 200 Ω resistive load over the frequency range of 20–20,000 Hz in order to deliver a required amount of power without exceeding its maximum voltage, current, and power ratings, whereas the speaker it has to drive might have an impedance of 8 Ω; hence transformer coupling is required. Transformer coupling is also employed where the load resistor must "float" referenced to the input voltage source both for dc and for ac signals.

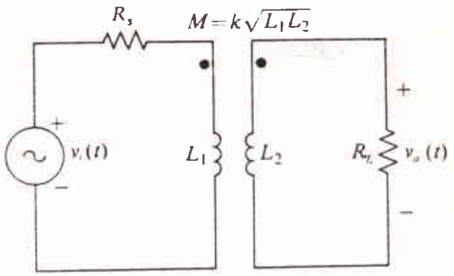


Fig. 2.1–1 Transformer coupling network.

Since our objective in this section is to gain familiarity with the basic operation of the transformer as a coupling element, we neglect second-order effects such as winding capacity and core nonlinearities in our analysis. In addition, we model the resistive losses in the transformer as small resistors in series with the input and output terminals. This model is quite reasonable where core loss is not excessive in comparison with winding loss, as is the case in most commercial broadband transformers.

The transformer model most useful for analyzing broadband coupling networks is shown in Fig. 2.1–2, where r_1 and r_2 represent the transformer loss. The equivalence of this model, as well as other possible models, and the original transformer is explored in the appendix and in the problems at the end of the chapter. The model explicitly indicates the cause of the loss of high- and low-frequency transmission. In particular, at low frequencies the impedance of $L_b = k^2L_1$ approaches zero and shunts to ground the signal path to R_L; and at high frequencies the impedance of $L_a = (1 - k^2)L_1$ approaches infinity and thus opens the signal path to R_L. However, if $L_b \gg L_a$ (or equivalently $k \approx 1$), a frequency range exists where ωL_b is large in comparison with the impedance it shunts, while ωL_a is small in comparison with the impedance in series with it. Over this range, which we call the midband (cf. Fig. 2.1–7), the inductances L_a and L_b may be approximated by short and open circuits; this yields the simplified model shown in Fig. 2.1–3 for the network of Fig. 2.1–1.

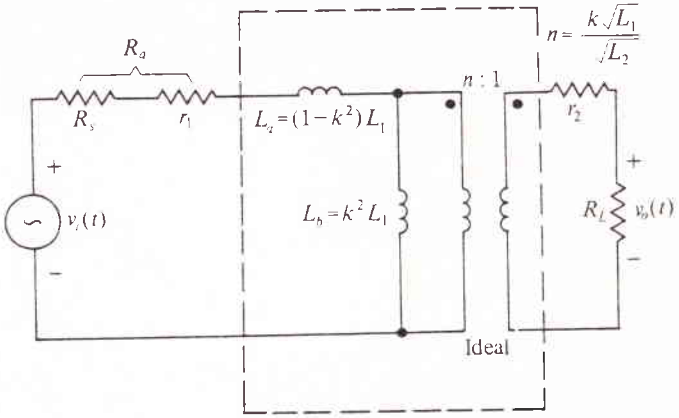


Fig. 2.1–2 Transformer model replacing transformer of Fig. 2.1–1.

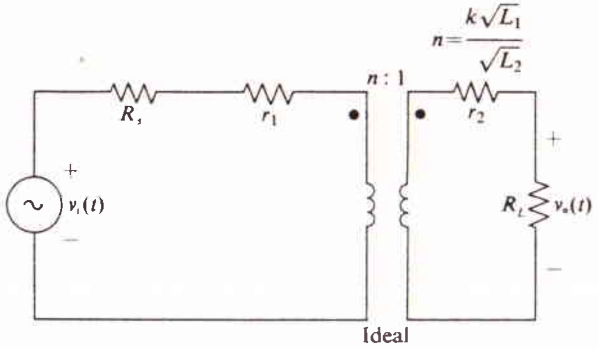


Fig. 2.1-3 Midband model for transformer coupling network.

With the aid of the midband model, we first observe that for the usual case where $r_2 \ll R_L$ and $r_1 \ll R_s$, an impedance of n^2R_L is presented to the driving source. Therefore, by choosing $n^2 = k^2L_1/L_2$ appropriately, we may obtain any load resistance required by the driving source. For $k \approx 1$ and L_1 and L_2 wound on the same core,

$$n^2 \approx \frac{L_1}{L_2} = \frac{N_1^2}{N_2^2},$$

where N_1 and N_2 are the numbers of turns in the windings of L_1 and L_2 respectively; hence in this case n may be related to the physical turns ratio of the transformer.

Second, we observe that in the midband n may be chosen to maximize the voltage across R_L for the case where $v_i(t)$, R_s, and R_L are fixed. Situations of this type arise when a transducer, such as a phonograph pickup, with a high source impedance (R_s) and a fixed developed signal (v_i) must be coupled in the midband to an amplifier with low input resistance (R_L). Writing the midband transfer function in the form

$$H_m = \frac{v_o(t)}{v_i(t)} = \frac{nR_L}{r_1 + R_s + n^2(r_2 + R_L)} \tag{2.1-1}$$

and equating dH_m/dn with zero, we obtain the value of n which maximizes H_m:

$$n_m = \sqrt{(R_s + r_1)/(R_L + r_2)}. \tag{2.1-2}$$

With this value of $n = n_m$, the midband transfer function is given by

$$H_m = \frac{R_L}{2\sqrt{(R_s + r_1)(R_L + r_2)}}. \tag{2.1-3}$$

The value of $n = n_m$ given in Eq. (2.1-2) is intuitively reasonable as the value which produces maximum signal to R_L, since it yields a resistance at the transformer input terminals which is equal to the source resistance $R_s + r_1$. Such a match ensures maximum power into the transformer and thus into R_L.

It is apparent from Eq. (2.1-3) that the existence of transformer loss reduces the signal available to R_L. To obtain a better measure of this signal attenuation, we assume that $v_i(t)$ is of the form $V_1 \cos \omega t$, where ω is some midband radian frequency, and that $n = n_m$, and then compute the ratio of the average power delivered to R_L to the average power delivered to R_L with $r_1 = r_2 = 0$. This ratio, which is a measure of efficiency η, is given by

$$\eta = \frac{P_L}{P_L|_{r_1=r_2=0}} = \frac{\dfrac{V_1^2}{2R_L}\dfrac{R_L^2}{4(R_s + r_1)(R_L + r_2)}}{\dfrac{V_1^2}{2R_L}\dfrac{R_L^2}{4R_sR_L}} = \frac{R_s}{R_s + r_1}\frac{R_L}{R_L + r_2}. \tag{2.1-4}$$

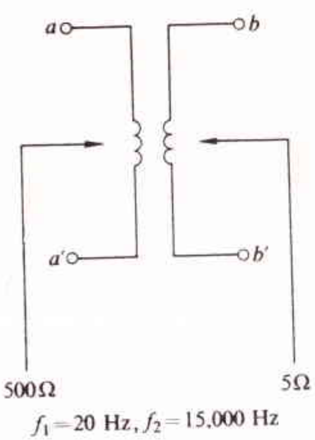


Fig. 2.1-4 Typical transformer specification.

Clearly then, unless $r_1 \ll R_s$ and $r_2 \ll R_L$, much of the available signal power is not supplied to R_L.

The transformer manufacturer usually indicates what minimum values of R_s and R_L ensure $R_L \gg r_2$ and $R_s \gg r_1$ in his specification of the turns ratio. Typically, the specification appears in the form shown in Fig. 2.1-4, which is interpreted to imply that if a 5 Ω resistor is connected across terminals b and b', 500 Ω is "seen" at terminals a and a' in the midband range extending from 20 Hz to 15,000 Hz; hence $n = \sqrt{500/5} = 10$. In addition, the fact that $R_L = 5\,\Omega$ and $R_s = 500\,\Omega$ ensures $R_L \gg r_2$ and $R_s \gg r_1$, usually to the extent that $\eta > 0.8$. Using smaller values of R_L does not alter the turns ratio n but does decrease the efficiency η and alter the midband frequency range.

To extend our analysis beyond the range of the midband, we obtain the transfer function $H(p) = V_o(p)/V_1(p)$ for the circuit of Fig. 2.1-2 in the form

$$H(p) = \frac{\dfrac{nR_L}{L_a}p}{p^2 + p\left(\dfrac{R_a}{L_a} + \dfrac{R_b}{L_b} + \dfrac{R_b}{L_a}\right) + \dfrac{R_aR_b}{L_aL_b}}, \tag{2.1-5}$$

where $R_a = R_s + r_1$, $R_b = n^2(R_L + r_2)$, $L_a = (1 - k^2)L_1$, and $L_b = k^2L_1$. Since all RL (or RC) networks have their poles on the negative real axis, the pole-zero

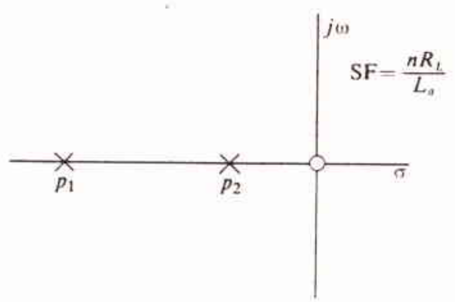


Fig. 2.1–5 Pole-zero diagram of $H(p)$.

diagram for $H(p)$ takes the form shown in Fig. 2.1–5, where p_1 and p_2 are the roots of the denominator of $H(p)$. In general, the expression for p_1 and p_2 is quite complicated; however, for the broadband transformer where $L_b \gg L_a$, and in turn $p_1 \gg p_2$, simplified approximate expressions for p_1 and p_2 may be obtained.

We note from Eq. (2.1–5) that the sum and product of the roots of the denominator of $H(p)$ are given by

$$p_1 + p_2 = -\left(\frac{R_a + R_b}{L_a} + \frac{R_b}{L_b}\right) \qquad \text{and} \qquad p_1 p_2 = \frac{R_a R_b}{L_a L_b}.$$

As L_b increases relative to L_a, $p_1 + p_2$ approaches a constant value while $p_1 p_2$ approaches zero. But $p_1 p_2$ can approach zero only if one of the poles, p_2 in this case, approaches the origin. Thus as L_b increases relative to L_a the larger pole is approximated by the sum of the poles; that is,

$$p_1 \approx p_1 + p_2 = -\left(\frac{R_a + R_b}{L_a} + \frac{R_b}{L_b}\right) \approx -\frac{R_a + R_b}{L_a} \equiv p_{10} \tag{2.1–6}$$

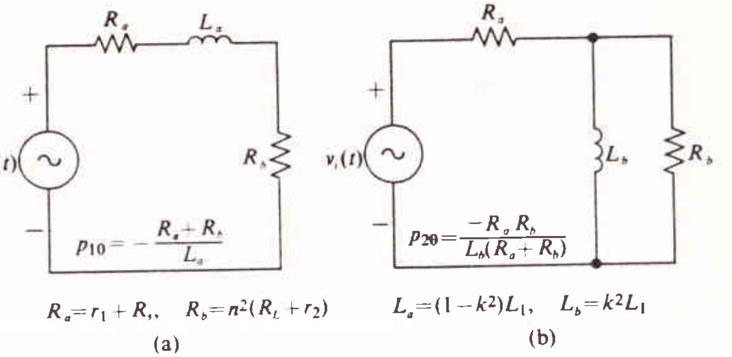
and

$$p_2 \approx \frac{p_1 p_2}{p_{10}} = \frac{-\dfrac{R_a R_b}{L_a L_b}}{\dfrac{R_a + R_b}{L_a}} = -\frac{R_a R_b}{R_a + R_b}\frac{1}{L_b} \equiv p_{20}. \tag{2.1–7}$$

As the reader can readily demonstrate numerically, if $R_a \approx R_b$ and $L_b > 10L_a$, the approximations of Eqs. (2.1–6) and (2.1–7) are accurate within 5%. In addition, if $L_b > 100L_a$ the approximations are valid within 1% for any ratio of R_a to R_b. It should be noted that p_{10} is the network pole obtained with L_b open-circuited and that p_{20} is the network pole obtained with L_a replaced by a short circuit. Figure 2.1–6 shows the two simplified single-pole circuits from which p_{10} and p_{20} may be obtained by inspection.

With the poles widely separated, $H(p)$ is given by

$$H(p) = \frac{\dfrac{nR_L}{L_a}p}{(p - p_{10})(p - p_{20})} = \frac{\dfrac{nR_L}{L_a}p}{\left(p + \dfrac{R_a + R_b}{L_a}\right)\left(p + \dfrac{1}{L_b}\dfrac{R_a R_b}{R_a + R_b}\right)}; \tag{2.1–8}$$


R_a
L_a
$+$
$v_i(t)$
$-$
R_b
$p_{10} = -\frac{R_a + R_b}{L_a}$
R_a
$+$
$v_i(t)$
$-$
L_b
R_b
$p_{20} = \frac{-R_a R_b}{L_b(R_a + R_b)}$

$R_a = r_1 + R_s$, $R_b = n^2(R_L + r_2)$

(a)

$L_a = (1 - k^2)L_1$, $L_b = k^2 L_1$

(b)

Fig. 2.1–6 Simplified circuits for determining p_{10} and p_{20}.

in addition,

$$20 \log |H(j\omega)| = 20 \log H_m + 20 \log |\omega/p_{20}| - 20 \log \sqrt{1 + (\omega/p_{10})^2} - 20 \log \sqrt{1 + (\omega/p_{20})^2}, \tag{2.1–9}$$

where H_m is the midband transfer function. The magnitude $20 \log |H(j\omega)|$ and a sketch of arg $H(j\omega)$ vs. ω are given in Fig. 2.1–7.

Since $|p_{10}| \gg |p_{20}|$, the corrections at each corner to the asymptotes do not interact; hence the range of frequencies over which $|H(j\omega)|$ has not decreased by more than 3 dB from its midband value is just the range between $|p_{10}|$ and $|p_{20}|$. This range is conventionally defined as the -3 dB bandwidth of the transformer coupling network. One should note also that as the poles become widely separated, arg $H(j\omega)$ approaches $+\pi/4$ at $\omega = |p_{20}|$ and $-\pi/4$ at $\omega = |p_{10}|$.

If the transformer coupling network is now excited by a step of voltage of the form $v_i(t) = V_1 u(t)$, then $V_i(p) = V_1/p$, $V_o(p) = V_1 H(p)/p$, and

$$v_o(t) = \mathscr{L}^{-1}\frac{V_1 H(p)}{p} = \frac{V_1 n R_L}{L_a|p_1 - p_2|}(e^{-|p_2|t} - e^{-|p_1|t})u(t), \tag{2.1–10}$$

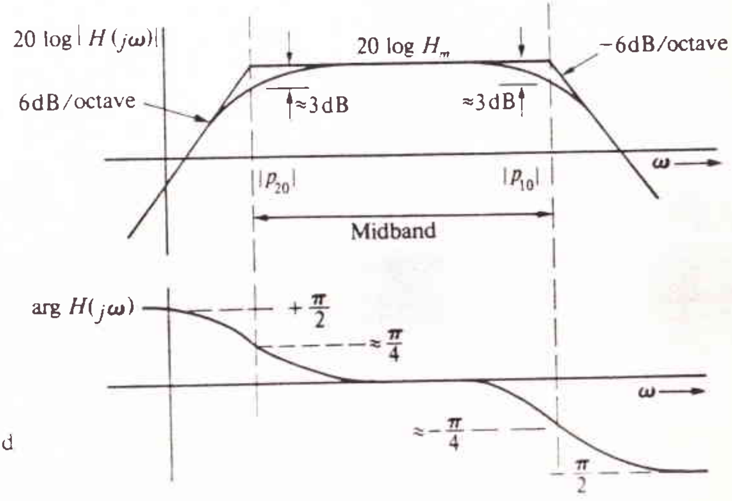


Fig. 2.1–7 Magnitude and phase vs. ω for transformer coupling network.

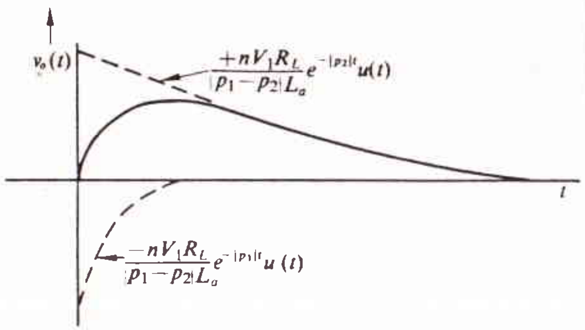


Fig 2.1–8 Step response of transformer coupling network.

where $\mathscr{L}^{-1}$ is the inverse Laplace transform operator. A plot of $v_o(t)$ vs. t is shown in Fig. 2.1–8. Note that the high-frequency pole at p_1 contributes to the deterioration of the leading edge of the output step, while the low-frequency pole at p_2 contributes to the decay of the output step to zero. This, of course, is an expected result, since a transformer does not transmit either the high-frequency components of the step which contribute to its leading edge or the low-frequency (dc) components which are required for a nonzero steady-state value.

If one makes the transformer broadband by designing $k \approx 1$ (for which case $p_1 \approx p_{10}$ and $p_2 \approx p_{20}$), the step response takes the form shown in Fig. 2.1–9. Since $|p_1| \gg |p_2|$, the time duration of $e^{-|p_1|t}$ becomes negligible when compared with $e^{-|p_2|t}$; thus the step response takes the approximate form

$$v_o(t) = \frac{V_1nR_L}{L_ap_{10}}e^{-|p_{20}|t}u(t) = \frac{V_1nR_L}{R_a + R_b}e^{-|p_{20}|t}u(t). \tag{2.1–11}$$

Note that this is the response obtainable from the simplified circuit of Fig. 2.1–6b. If the fine structure of the leading edge of the step response is required, an expanded time scale about the origin must be employed. On such a scale $e^{-|p_2|t}$ remains essentially constant at unity; thus the leading edge of the step response takes the approximate form

$$v_o(t) \approx \frac{V_1nR_L}{R_a + R_b}(1 - e^{-|p_{10}|t})u(t), \tag{2.1–12}$$

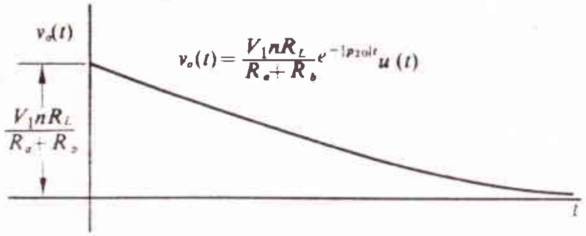


Fig. 2.1–9 Step response of broadband transformer network.

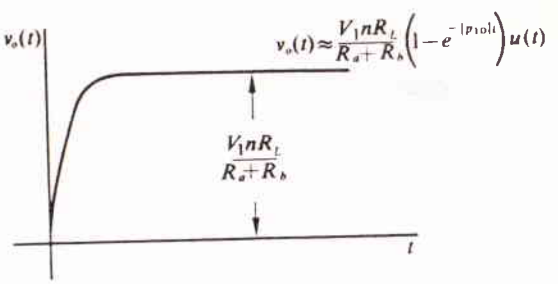


Fig. 2.1–10 Expansion of leading edge of Fig. 2.1–9.

which is illustrated in Fig. 2.1–10 and which is exactly the step response obtained from the simplified circuit of Fig. 2.1–6(a). Consequently, for the case of widely separated poles, both the step response and the frequency response may be obtained from simplified single-pole circuits.

As an application of the above analysis, let us specify the parameters of a transformer which matches 5 Ω (R_L) to 500 Ω (R_s) in the midband extending from 50 Hz to 5000 Hz ($|p_1| = 3.14 \times 10^4$ rad/sec, $|p_2| = 3.14 \times 10^2$ rad/sec). With these values the poles are widely separated; hence

$$L_a = (1 - k^2)L_1 = \frac{R_s + r_1 + n^2(R_L + r_2)}{|p_1|},$$

$$L_b = k^2L_1 = \frac{(R_s + r_1)(R_L + r_2)n^2}{[R_s + r_1 + n^2(R_L + r_2)]|p_2|}.$$

If the transformer is to have reasonable efficiency, we must have $R_s \gg r_1$ and $R_L \gg r_2$. In addition, for a match, we need $n^2R_L = R_s$; therefore,

$$(1 - k^2)L_1 \approx \frac{1000}{3.14 \times 10^4} = 31.8 \text{ mH} \qquad \text{and} \qquad k^2L_1 \approx \frac{250}{314} = 800 \text{ mH},$$

from which we obtain $k = 0.962$, $L_1 = 832$ mH, and $L_2 = k^2L_1/n^2 = 800$ mH/100 = 8 mH. All of these values are readily obtained in practice.

Before terminating our discussion of broadband transformer networks, let us consider the effect of loading the transformer in a way different from that specified by the manufacturer. Clearly if both R_s and R_L are increased while their ratio remains constant, the efficiency increases and $|p_{10}|$ and $|p_{20}|$ increase in direct proportion to R_s or R_L (if we assume $r_1 \ll R_s$ and $r_2 \ll R_L$); hence the transformer passband is shifted up in frequency. If R_L and R_s are decreased, the opposite effect results. To determine the effect on $|p_{10}|$ and $|p_{20}|$ of varying R_L relative to R_s, we plot

$$20 \log |p_{10}| = 20 \log \frac{R_b}{L_a}\left(1 + \frac{R_a}{R_b}\right)$$

and

$$20 \log |p_{20}| = 20 \log \frac{R_b}{2L_b}\left(\frac{2}{1 + R_b/R_a}\right)$$

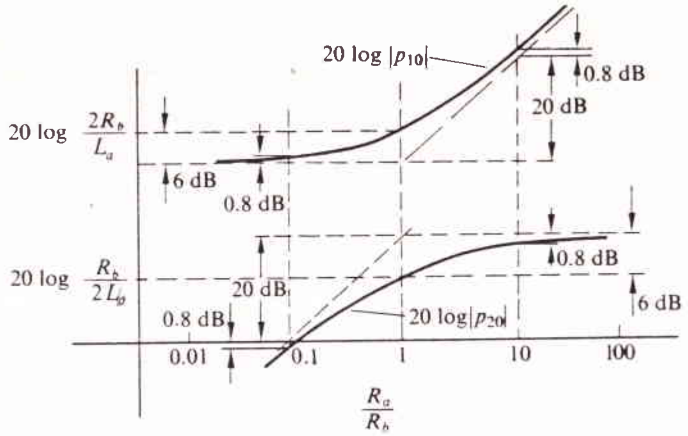


Fig. 2.1–11 Plot of 20 log $|p_{10}|$ and 20 log $|p_{20}|$ vs. R_a/R_b.

vs. R_a/R_b as shown in Fig. 2.1–11, where R/L is expressed as a dimensionless quantity in some convenient system of units. As $R_a = R_s + r_1$ is increased relative to $R_b = n^2(R_L + r_2)$, we observe that the high-frequency pole increases without bound while the low-frequency pole approaches a constant value. Consequently, we should expect a current source drive ($R_s = \infty$) to produce no high-frequency break point regardless of the value of k. This effect is not observed in practice, however, since an additional finite break point is produced by the physical winding capacity, which we have neglected.

As R_a is decreased relative to R_b, we observe that the high-frequency pole approaches a constant value while the low-frequency pole approaches zero. In the limit as $R_a \to 0$ we should expect the transformer to pass dc; however, this effect is also not observed in practice because of the nonzero resistance of the driving source and the existence of r_1. Although the above analysis is performed for a transformer coupling network, the same results are obtained foı any broadband coupling network having a zero at the origin and two poles on the negative real axis. In particular, a network of the form shown in Fig. 2.1–12, consisting of a coupling capacitor between the driver and the load with some stray capacity to ground, has a similar step and frequency response.

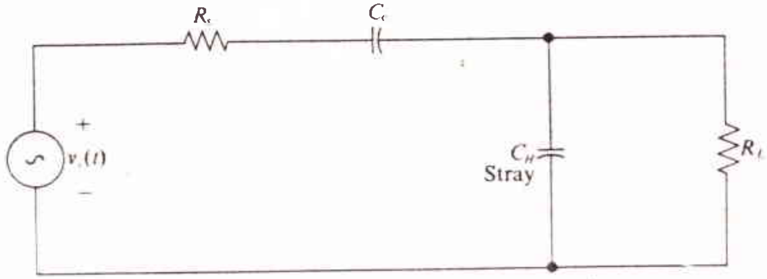


Fig. 2.1–12 Capacitive coupling network.

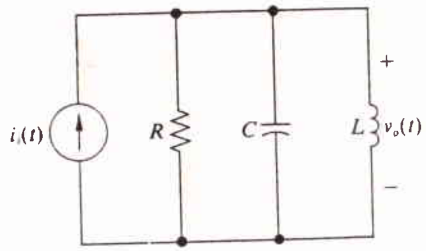


Fig. 2.2–1 Parallel *RLC* circuit.

In the sections which follow we shall shift our emphasis from the broadband application of the transformer to its application in narrowband circuits. We shall begin by developing the properties of simple narrowband circuits; we shall then extend these concepts to narrowband transformer-coupled networks.

2.2 PARALLEL *RLC* CIRCUIT

Before beginning a discussion of narrowband transformer networks, we shall review some of the properties of simple narrowband resonant circuits, which are essential building blocks of the more complicated transformer networks. We shall begin by considering the simple parallel *RLC* circuit shown in Fig. 2.2–1. If we drive the circuit with the current $i_i(t)$ and define $v_o(t)$ as the output voltage, the transfer function (in this case, input impedance) takes the form

$$Z_{11}(p) = \frac{V_o(p)}{I_i(p)} = \frac{p\dfrac{1}{C}}{p^2 + \dfrac{p}{RC} + \dfrac{1}{LC}} = \frac{p\dfrac{1}{C}}{(p - p_1)(p - p_2)}, \tag{2.2–1}$$

where

$$p_{1,2} = -\frac{1}{2RC} \pm \sqrt{\left(\frac{1}{2RC}\right)^2 - \frac{1}{LC}}.$$

The poles, p_1 and p_2, may be real or a complex conjugate pair. We consider these two cases separately.

For the case where $(1/2RC)^2 \geq 1/LC$ (or equivalently $R \leq \omega_0 L/2 = 1/2\omega_0 C$, where $\omega_0 = 1/\sqrt{LC}$), p_1 and p_2 lie on the negative real axis as shown in Fig. 2.2–2. This pole-zero diagram, which is analogous to the one for the broadband transformer network of Section 2.1, indicates that for values of R which are small compared with

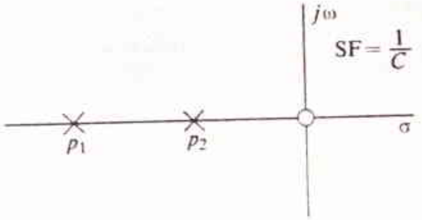
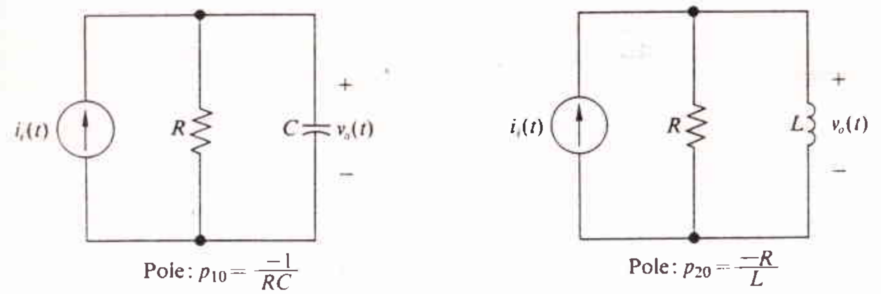


Fig. 2.2–2 Pole-zero diagram of $Z_{11}(p)$ with $R < \omega_0 L/2$.



Pole: $p_{10} = \frac{-1}{RC}$ Pole: $p_{20} = \frac{-R}{L}$

Fig. 2.2–3 Simplified circuits for finding p_{10} and p_{20}.

$\omega_0 L/2$, the parallel RLC network functions as a broadband network whose midband frequency range extends from $|p_2|$ to $|p_1|$. Physically the midband may be interpreted as that range of frequencies over which the impedance of both the inductance L and capacitance C is large, and therefore negligible, compared with R; hence $Z_{11_m} = R$. At low and high frequencies the impedances of L and C respectively approach zero, thus shunting $i_i(t)$ to ground.

For wide pole separation the approximate values of p_1 and p_2 may be determined by inspection. In particular, for $p_1 \gg p_2$, p_1 is approximated by the sum of the roots of the denominator of $Z_{11}(p)$, and p_2 is approximated by the product of the roots divided by the sum of the roots, that is,

$$p_1 \approx -\frac{1}{RC} \equiv p_{10}; \qquad p_2 = \frac{-1/LC}{1/RC} = \frac{R}{L} \equiv p_{20}. \tag{2.2–2}$$

Physically $|p_{20}|$ is the radian frequency at which the impedance of L is equal to R, and $|p_{10}|$ is the radian frequency at which the impedance of C is equal to R. Alternatively p_{10} and p_{20} are the poles of the simplified circuits shown in Fig. 2.2–3. Since the step and frequency responses of the broadband parallel RLC network in terms of p_1 and p_2 are identical in form with the corresponding responses of the broadband transformer, we do not present them here.

For the case where $(1/2RC)^2 < 1/LC$ or $R > \omega_0 L/2$, the expression for the poles takes the form

$$p_{1,2} = -\alpha \pm j\sqrt{\omega_0^2 - \alpha^2} = -\alpha \pm j\beta, \tag{2.2–3}$$

where $\omega_0 = 1/\sqrt{LC}$, $\alpha = 1/2RC$, and $\beta = \sqrt{\omega_0^2 - \alpha^2}$. The pole-zero diagram for this case is shown in Fig. 2.2–4. It is apparent that the distance of the poles from the origin is given by $\alpha^2 + \beta^2 = \omega_0^2$; hence as α is increased by increasing the loading on the circuit (decreasing R), the poles move into the left half-plane along the semicircular trajectory of radius ω_0 until they meet on the real axis for $\alpha = \omega_0$.

Physically ω_0 is that radian frequency at which the impedances of the inductor and capacitor are equal in magnitude and opposite in phase, and thus produce an open circuit when combined in parallel. This frequency, at which the parallel RLC circuit appears as a pure resistor R, is called the resonant frequency of the circuit.

UNIVERSIDAD DE
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION
Y BIBLIOTECA
MONTEVIDEO - URUGUAY

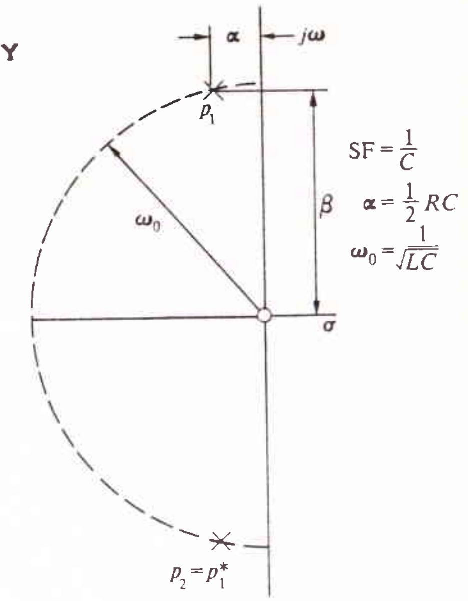


Fig. 2.2–4 Pole-zero diagram of $Z_{11}(p)$ with $R > \omega_0 L/2$.

To obtain the sinusoidal steady-state frequency response when $R > \omega_0 L/2$, we express $Z_{11}(j\omega)$ as

$$Z_{11}(j\omega) = \frac{j\omega/C}{\omega_0^2 - \omega^2 + 2j\omega\alpha} = \frac{R}{1 + jQ_T\dfrac{\omega^2 - \omega_0^2}{\omega\omega_0}}, \tag{2.2–4}$$

where $Q_T = \omega_0/2\alpha = \omega_0 RC = R/\omega_0 L$. Equation (2.2–4) provides the means of obtaining an exact plot of $|Z_{11}(j\omega)|$ and $\arg Z_{11}(j\omega)$ vs. ω. On the other hand, a quick sketch of $|Z_{11}(j\omega)|$ and $\arg Z_{11}(j\omega)$ may be obtained from the value of $Z_{11}(j\omega_0)$ and the asymptotic values of $Z_{11}(j\omega)$ as ω approaches zero and infinity. The asymptotic values vs. ω may be obtained by plotting the resistance R, the magnitude of the capacitive reactance $|X_C(\omega)| = 1/\omega C$, and the magnitude of the inductive reactance $|X_L(\omega)| = \omega L$ on the same set of coordinates vs. ω as shown in Fig. 2.2–5. Clearly for $\omega = \omega_0$ the reactances of the inductor and capacitor cancel each other and yield $|Z_{11}(j\omega)| = R$ and, of course, $\arg Z_{11}(j\omega) = 0$. As ω is decreased below ω_0, the small reactance of the inductor rapidly dominates the parallel RLC network, causing $|Z_{11}(j\omega)|$ to approach ωL and $\arg Z_{11}(j\omega)$ to approach $\pi/2$. Similarly, as ω is increased above ω_0, the small reactance of the capacitor rapidly dominates the parallel RLC network, causing $|Z_{11}(j\omega)|$ to approach $1/\omega C$ and $\arg Z_{11}(j\omega)$ to approach $-\pi/2$, as shown in Fig. 2.2–5.

In addition, the two "half-power" frequencies ω_1 and ω_2, at which $|Z_{11}(j\omega)| = R/\sqrt{2}$ and $\arg Z_{11}(j\omega) = +\pi/4$ and $-\pi/4$ respectively, may be found by equating the imaginary term in the denominator of Eq. (2.2–4) to ∓ 1 and solving for ω:

$$Q_T\frac{\omega^2 - \omega_0^2}{\omega\omega_0} = \begin{cases} -1, & \omega = \omega_1, \\ +1, & \omega = \omega_2. \end{cases} \tag{2.2–5}$$

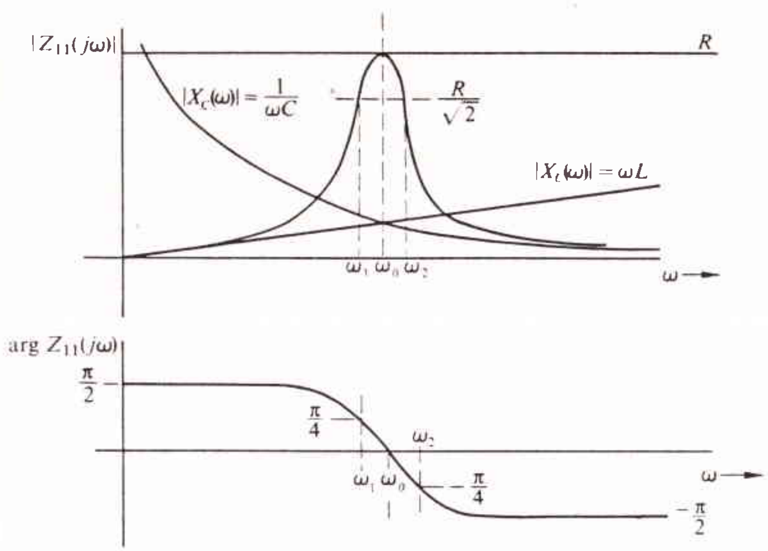


Fig. 2.2–5 Plot of $|Z_{11}(j\omega)|$ and $\arg |Z_{11}(j\omega)$ vs. ω.

Direct solution of Eq. (2.2–5) yields the following relationships between ω_1 and ω_2:

$$\omega_2 - \omega_1 = 2\alpha \qquad \text{and} \qquad \omega_1\omega_2 = \omega_0^2.$$

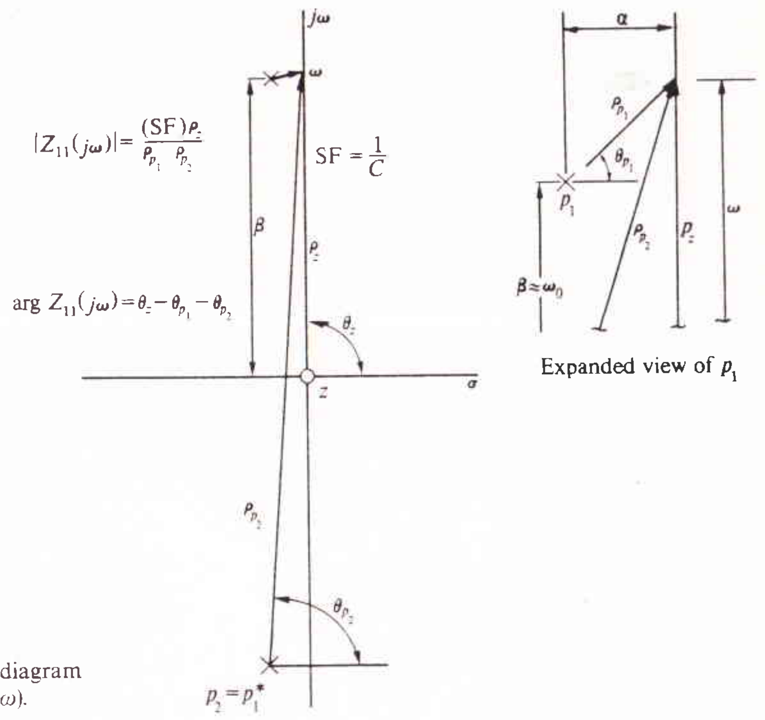
However, since the difference between ω_2 and ω_1 is the -3 dB bandwidth (BW) of the parallel RLC circuit, we have BW $= 2\alpha$, or equivalently

$$\frac{\text{BW}}{\omega_0} = \frac{2\alpha}{\omega_0} = \frac{1}{Q_T}. \tag{2.2–6}$$

It is interesting at this point to note the significance of the parameter Q_T, which is referred to as the network "que." First of all, as Q_T increases, the -3 dB bandwidth of the network decreases relative to its resonant frequency. Clearly, then, if the parallel RLC circuit is used to transmit a modulated carrier whose spectrum occupies a fixed band of frequencies about ω_0, Q_T must not be increased to a point where the modulation is distorted. Second, increasing Q_T increases the ratio of R to $\omega_0 L = 1/\omega_0 C$ (note that $Q_T = R/\omega_0 L = R\omega_0 C$). Hence the parallel RLC network greatly attenuates frequencies in the vicinity of $2\omega_0$ (where $|Z_{11}(j\omega)| \rightarrow 1/\omega C$) relative to frequencies in the vicinity of ω_0 (where $Z_{11}(j\omega) \approx R$). Consequently, if the purpose of the parallel RLC circuit is to extract the fundamental component of a periodic waveform, a Q_T as high as possible is desired. Third, as Q_T increases, the poles of $Z_{11}(p)$ approach the imaginary axis in the complex p-plane and cause any transients induced in the network to become more and more oscillatory.

If we wish, we may interpret Q_T in still another way. For the case where $i_i(t) = I\cos\omega_0 t$, which results in $v_o(t) = IR\cos\omega_0 t = V_1\cos\omega_0 t$, we have

$$\frac{2\pi\,(\text{peak energy stored})}{\text{energy dissipated per cycle}} = \frac{2\pi(\frac{1}{2}CV_1^2)}{(2\pi/\omega_0)(V_1^2/2R)} = \omega_0 CR = Q_T. \tag{2.2–7}$$



Expanded view of p_1

Fig. 2.2–6 Phasor diagram for evaluating $Z_{11}(j\omega)$.

This interpretation of Q_T in terms of stored and dissipated energy is useful in determining the Q of circuits with more than one source of dissipation.

In cases where Q_T is high, or equivalently where the poles of $Z_{11}(p)$ are close to the imaginary axis, a graphical approximation permits us to obtain a greatly simplified expression for $Z_{11}(j\omega)$. If we consider the phasor diagram shown in Fig. 2.2–6, we observe that for $\omega > 0$ the range of frequencies over which $|Z_{11}(j\omega)| = (SF)\rho_z/\rho_{p_1}\rho_{p_2}$ is significantly different from zero is that range in the vicinity of p_1 where ρ_{p_1} becomes small. As α is decreased relative to ω_0, this frequency range decreases to the point where the phasors drawn from z and $p_2 = p_1^*$ remain essentially constant over the entire range, with the magnitudes and angles given by

$$\rho_z \approx \omega_0, \qquad \theta_z = \pi/2, \qquad \rho_{p_2} \approx 2\omega_0, \qquad \theta_{p_2} \approx \pi/2.$$

In addition, for $\alpha \ll \omega_0$,

$$\beta = \sqrt{\omega_0^2 - \alpha^2} \approx \omega_0.$$

Consequently, for high values of Q_T, $Z(j\omega)$ may be closely approximated by

$$Z(j\omega) \approx \frac{\omega_0}{C(2\omega_0)\rho_{p_1}} \exp j\left(\frac{\pi}{2} - \frac{\pi}{2} - \theta_{p_1}\right) \tag{2.2–8}$$

$$= \frac{\exp(-j\theta_{p_1})}{2C\rho_{p_1}}.$$

Recognizing that $\rho_{p_1} \exp j\theta_{p_1} = \alpha + j(\omega - \beta) \approx \alpha + j(\omega - \omega_0)$, we may rewrite Eq. (2.2-8) in the form

$$Z_{11}(j\omega) = \frac{1}{2\alpha C\left(1 + j\dfrac{\omega - \omega_0}{\alpha}\right)} = \frac{R}{1 + j\dfrac{\omega - \omega_0}{\alpha}} \qquad \omega > 0. \tag{2.2-9}$$

A similar form may be obtained for $\omega < 0$.

It is apparent that $|Z_{11}(j\omega)|$ given by Eq. (2.2-9) is symmetric about ω_0, with "half-power" frequencies ω_1 and ω_2 given by

$$\omega_{1,2} = \omega_0 \pm \alpha.$$

It is interesting that for high values of Q_T, ω_0 approaches the arithmetic rather than the geometric mean of ω_1 and ω_2, while the -3 dB bandwidth remains unchanged at $\omega_2 - \omega_1 = 2\alpha$.
For $\omega = \omega_1$ and for $\omega = \omega_2$, ρ_{z_1} has increased from its minimum value of α to $\alpha\sqrt{2}$, thus causing $|Z_{11}(j\omega)|$ to decrease from its maximum value by a factor of $1/\sqrt{2}$.

Although we have indicated that the simplified form of $Z_{11}(j\omega)$ given by Eq. (2.2-9) is a good approximation for $Z_{11}(j\omega)$ when Q_T is high, we have not yet determined what range of Q_T may be considered high. To do this we manipulate the exact expression for $Z_{11}(j\omega)$ given by Eq. (2.2-4) into the form

$$Z_{11}(j\omega) = \frac{R}{1 + j\Omega\dfrac{\Omega + 4Q_T}{2\Omega + 4Q_T}}, \tag{2.2-10}$$

where $\Omega = (\omega - \omega_0)/\alpha$. Here we see explicitly that as $Q_T \to \infty$ Eq. (2.2-10) reduces to Eq. (2.2-9). In addition, if we plot $|Z_{11}(j\omega)|/R$ and $\arg Z_{11}(j\omega)$ vs. Ω as shown in Figs. 2.2-7 and 2.2-8, we observe exceptionally close agreement between the curves obtained for $Q_T = 10$ and $Q_T = \infty$, particularly in the vicinity of $\Omega = 0$, or equivalently $\omega = \omega_0$. Since the curves for $Q_T = \infty$ correspond to the simplified expression for $Z_{11}(j\omega)$ given by Eq. (2.2-9), we may clearly use this approximation for $Z_{11}(j\omega)$ with confidence for $Q_T > 10$ and as a "ballpark" approximation for Q_T as low as 5.

To obtain the response of the parallel resonant circuit driven by an impulse of current, we obtain the inverse Laplace transform of $Z_{11}(p)$ or directly evaluate the impulse response by other methods. In either case, we obtain

$$z_{11}(t) = \mathscr{L}^{-1}[Z_{11}(p)] = \frac{\omega_0}{C\beta}e^{-\alpha t}\cos\left(\beta t + \tan^{-1}\frac{\alpha}{\beta}\right)u(t). \tag{2.2-11}$$

For the case where $Q_T \geq 10$, Eq. (2.2-11) reduces to the simplified form

$$z_{11}(t) \approx \frac{1}{C}e^{-\alpha t}(\cos\omega_0 t)u(t), \tag{2.2-12}$$

since $\beta = \omega_0\sqrt{1 - 1/4Q_T^2} \approx \omega_0$ and $\alpha/\beta = 1/\sqrt{4Q_T^2 - 1} \approx 0$.

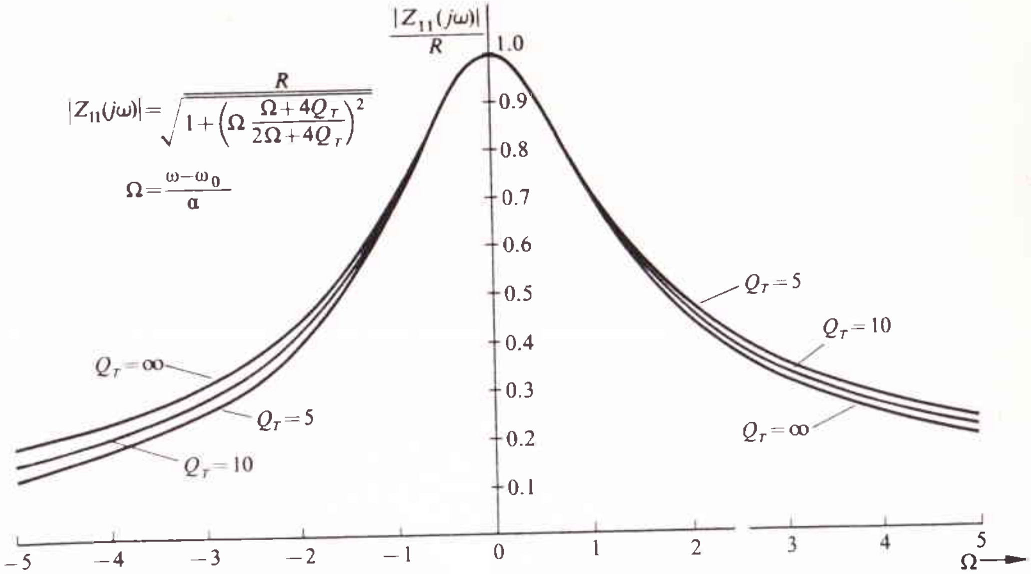


Fig. 2.2-7 Plot of $|Z_{11}(j\omega)|/R$ vs. $\Omega = (\omega - \omega_0)/\alpha$.

2.3 PARALLEL *LC* CIRCUIT WITH SERIES LOSS

In this section we shall consider the parallel resonant circuit with loss in series with the inductor or capacitor. We shall show that when the series loss is small, as is most often the case in communication systems, the circuit behaves in the same fashion as the parallel *RLC* network, and thus the loss may be modeled by an equivalent parallel resistor. Such a model permits the combination of circuit loss appearing at several points in the parallel *LC* circuit into a single parallel resistor for which the expressions derived in Section 2.2 may be applied directly.

We begin our analysis with the circuit of Fig. 2.3-1, in which the loss in the form of r appears in series with the inductor. We shall then generalize to the case where loss appears in series with both the inductor and capacitor and also in parallel with the combination. For the circuit of Fig. 2.3-1, the transfer function (or input impedance) relating $v_o(t)$ and $i_i(t)$ is given by

$$Z_{11}(p) = \frac{V_o(p)}{I_i(p)} = \frac{(1/C)(p + r/L)}{(p - p_1)(p - p_2)}, \tag{2.3-1}$$

where $p_{1,2} = (-r/2L) \pm \sqrt{(r/2L)^2 - 1/LC}$. We see that the poles of $Z_{11}(p)$ may be either real or complex, depending on the relationship of the parameters. Since the analysis of the circuit with real poles is quite similar to the real-pole analysis performed in Section 2.2, we restrict our attention to the case where $1/LC > (r/2L)^2$ (or equivalently $r < 2\omega_0 L = 2/\omega_0 C$, where $\omega_0 = 1/\sqrt{LC}$). For this case the poles may be written in the form

$$p_{1,2} = -\alpha \pm j\sqrt{\omega_0^2 - \alpha^2} = -\alpha \pm j\beta, \tag{2.3-2}$$

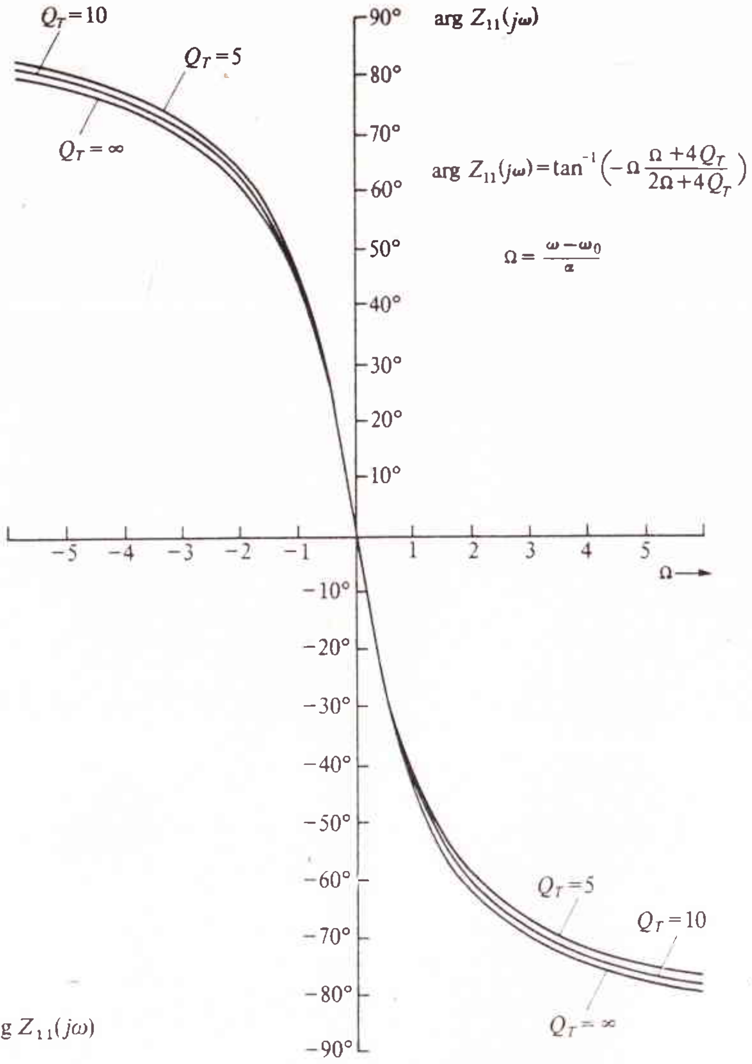


Fig. 2.2–8 Plot of arg $Z_{11}(j\omega)$ vs. $\Omega = (\omega - \omega_0)/\alpha$.

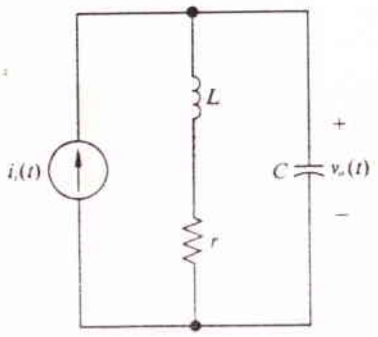


Fig. 2.3–1 Parallel *LC* circuit with loss in series with the inductor.

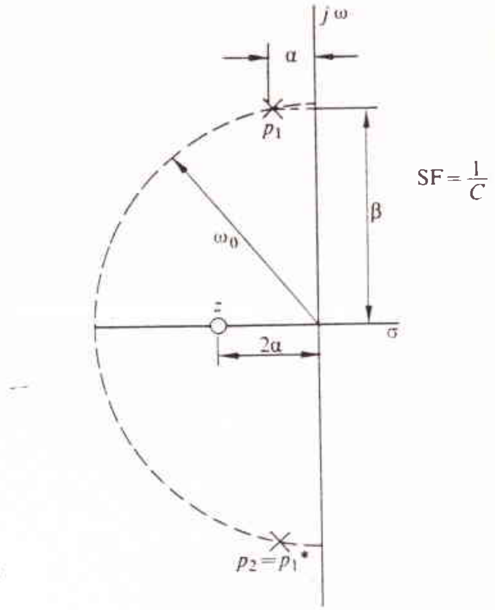


Fig. 2.3–2 Pole-zero diagram of $Z_{11}(p)$ for $r < 2\omega_0 L$.

where

$$\alpha = \frac{r}{2L}, \qquad \omega_0 = \frac{1}{\sqrt{LC}}, \qquad \text{and} \qquad \beta = \sqrt{\omega_0^2 - \alpha^2}.$$

The corresponding pole-zero diagram for $Z_{11}(p)$ is shown in Fig. 2.3–2. Here again we notice that as α is increased by increasing r, the poles move into the left half-plane along a semicircular trajectory of radius ω_0, where ω_0 is the frequency at which the magnitudes of the impedances of the inductor and capacitor are equal.

Figure 2.3–3 shows a sketch of $|Z_{11}(j\omega)|$ and arg $Z_{11}(j\omega)$ obtained graphically from the pole-zero plot of Fig. 2.3–2. Even this rough sketch indicates that, unlike the parallel *RLC* circuit, the circuit with loss in series with the inductor does not experience either zero phase shift or maximum amplitude of $|Z_{11}(j\omega)|$ at the frequency ω_0. For this reason two distinct resonant frequencies are specified for this circuit: (1) the amplitude resonant frequency at which the maximum value of $|Z_{11}(j\omega)|$ occurs and (2) the phase resonant frequency at which $Z_{11}(j\omega)$ appears purely resistive. Fortunately, as the poles p_1 and p_2 approach the imaginary axis, both resonant frequencies converge to ω_0, which is sometimes defined as still a third resonant frequency.

In addition, no convenient expression exists for either the maximum value of $|Z_{11}(j\omega)|$ or the −3 dB bandwidth of the circuit in the general case. However, for the case where $\alpha \ll \omega_0$ or Q_L† is high, where Q_L is defined as

$$Q_L \equiv \frac{\omega_0}{2\alpha} = \frac{\omega_0 L}{r} = \frac{1}{\omega_0 rC}. \tag{2.3–3}$$

certain simplifying approximations in the expressions for $Z_{11}(j\omega)$ may be made.

---

† We shall reserve Q_T for "que" in parallel *RLC* circuits.

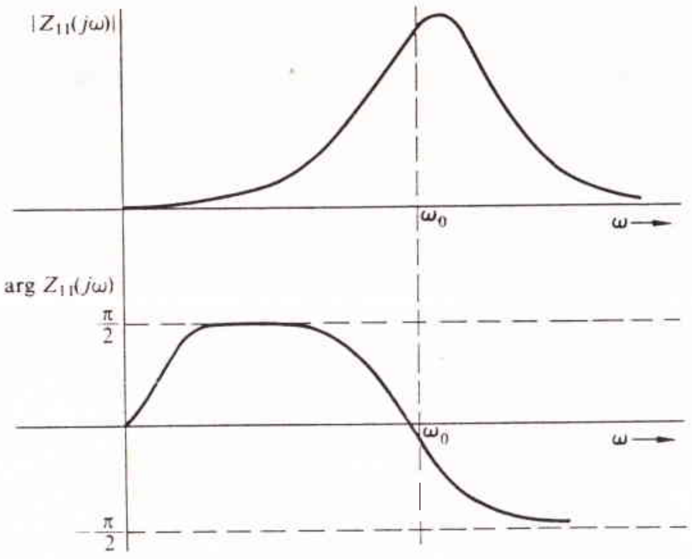


Fig. 2.3–3 Sketch of magnitude and phase of $Z_{11}(j\omega)$ vs. ω.

Specifically, with the aid of Fig. 2.3–4, we write $Z_{11}(j\omega)$ in the form

$$Z_{11}(j\omega) = \frac{(\mathrm{SF})\rho_z}{\rho_{p_1}\rho_{p_2}} \exp j(\theta_z - \theta_{p_1} - \theta_{p_2}) \tag{2.3–4}$$

and observe that, for frequencies in the vicinity of p_1 (the only range for $\omega > 0$ where $Z_{11}(j\omega)$ is significantly different from zero), the following approximations are valid:

$$\rho_z \approx \omega_0, \qquad \theta_z \approx \pi/2, \qquad \rho_{p_2} \approx 2\omega_0, \qquad \theta_{p_2} \approx \pi/2,$$

$$\beta = \sqrt{\omega_0^2 - \alpha^2} \approx \omega_0.$$

Thus Eq. (2.3–4) may be written in the form

$$Z_{11}(j\omega) \approx \frac{1}{2C\rho_{p_1}} \exp -j\theta_{p_1} = \frac{1}{2C\alpha\left(1 + j\dfrac{\omega - \omega_0}{\alpha}\right)}, \qquad \omega > 0, \tag{2.3–5}$$

which is identical in form to Eq. (2.2–9). If, in addition, we observe that

$$\frac{1}{2\alpha C} = \frac{\omega_0}{2\alpha}\frac{1}{\omega_0 C} = Q_L\omega_0 L = Q_L^2 r \equiv R_{\mathrm{eq}}, \tag{2.3–6}$$

then Eq. (2.3–5) may be rewritten in the form

$$Z_{11}(j\omega) = \frac{R_{\mathrm{eq}}}{1 + j\dfrac{\omega - \omega_0}{\alpha}}, \qquad \omega > 0. \tag{2.3–7}$$

which is exactly the expression for the impedance of a high-Q_T parallel RLC circuit with $R = R_{\mathrm{eq}} = Q_L^2 r$. Hence, for high values of Q_L [specifically, $Q_L > 10$ ensures the accuracy of Eq. (2.3–7) within a few percent], the circuit of Fig. 2.3–1 may be modeled by the circuit shown in Fig. 2.3–5. With this equivalence in circuits, we note from the results of Section 2.2 that $|Z_{11}(j\omega)|_{\max} = R_{\mathrm{eq}}$, $\mathrm{BW} = 2\alpha = r/L$, $\mathrm{BW}/\omega_0 = 1/Q_L$.

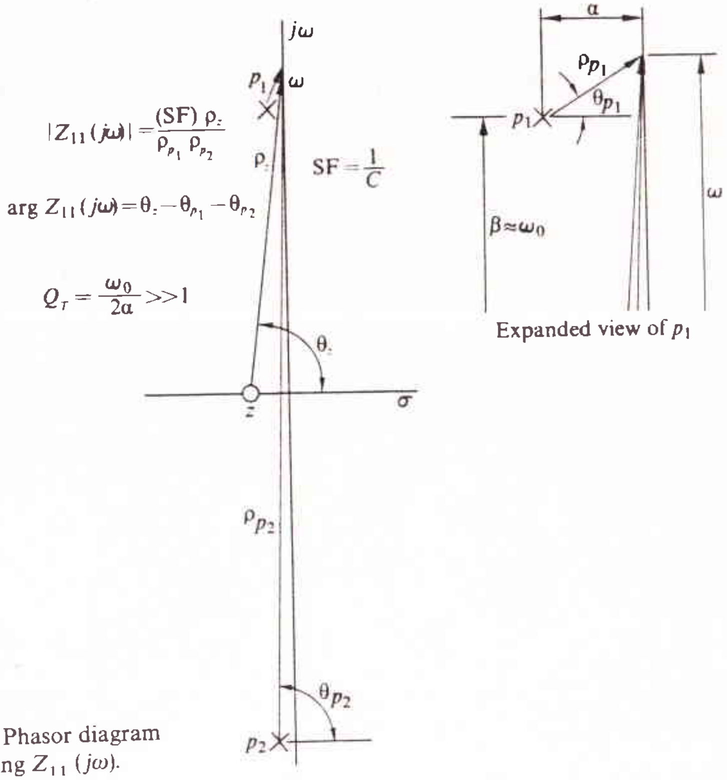
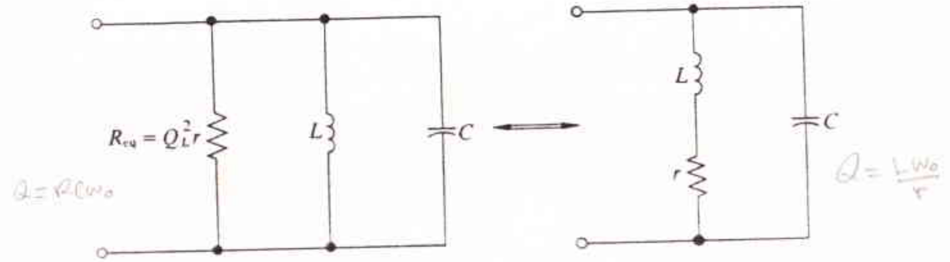

$j\omega$
p_1
ω
$|Z_{11}(j\omega)| = \frac{(\mathrm{SF})\,\rho_z}{\rho_{p_1}\,\rho_{p_2}}$
ρ_z
$\mathrm{SF} = \frac{1}{C}$
$\arg Z_{11}(j\omega) = \theta_z - \theta_{p_1} - \theta_{p_2}$
$Q_T = \frac{\omega_0}{2\alpha} >> 1$
θ_z
z
σ
ρ_{p_2}
θ_{p_2}
p_2
α
ρ_{p_1}
θ_{p_1}
p_1
$\beta \approx \omega_0$
ω
Expanded view of p_1

Fig. 2.3–4 Phasor diagram for evaluating $Z_{11}(j\omega)$.

In terms of the pole-zero diagram of $Z_{11}(p)$ shown in Fig. 2.3–2, the effect of replacing the series loss by a parallel loss is the movement of the zero at -2α to the origin. This movement has little effect on $Z_{11}(p)$ for small values of α, which exist when Q_L is high.



$$Q_L = \frac{\omega_0 L}{r} > 10$$

Fig. 2.3–5 Equivalence of two parallel resonant circuits.

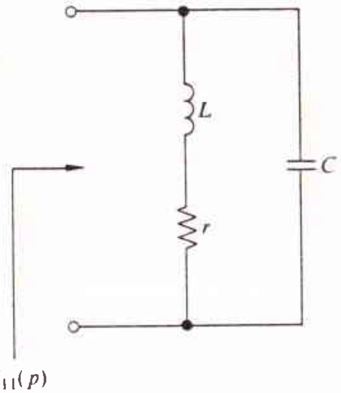


Figure 2.3–6

In practice, the quickest way to convert series loss into parallel loss is to evaluate $Q_L = \omega_0 L/r$, determine whether $Q_L > 10$, and then form $R_{eq} = Q_L^2 r$. Note that if $Q_L > 10$ and if there are no other losses in the network, then $Q_L = Q_T$. Clearly $Q_T = R_{eq}/\omega_0 L = Q_L$. If other losses are present in the circuit, R_{eq} must be combined with them to determine Q_T; thus $Q_L \neq Q_T$.

Example 2.3–1 Determine values for r, L, and C for the circuit shown in Fig. 2.3–6 such that $Z_{11}(j\omega)$ peaks to a value of 1000 Ω at $\omega_0 = 10^7$ rad/sec ($f = 1.6$ MHz) with a bandwidth of 5×10^5 rad/sec.

Solution. Since the bandwidth is narrow compared with ω_0, we have a high-Q circuit for which

$$Q_L = \frac{\omega_0}{2\alpha} = \frac{\omega_0}{\text{BW}} = 20.$$

In addition,

$$1000\ \Omega = |Z(j\omega_0)| \approx Q_L^2 r;$$

hence $r = 2.5\ \Omega$. From the relationship

$$\frac{r}{L} = 2\alpha \approx \text{BW}$$

we obtain

$$L = \frac{2.5\ \Omega}{5 \times 10^5\ \text{rad/sec}} = 5\ \mu\text{H},$$

and finally

$$C = 1/\omega_0^2 L = 2000\ \text{pF}.$$

If the loss in a parallel resonant circuit appears in series with the capacitor, as shown in Fig. 2.3–7, the input impedance is given by

$$Z_{11}(p) = \frac{V_o(p)}{I_i(p)} = \frac{rp(p + 1/rC)}{p^2 + p(r/L) + 1/LC}. \tag{2.3–8}$$

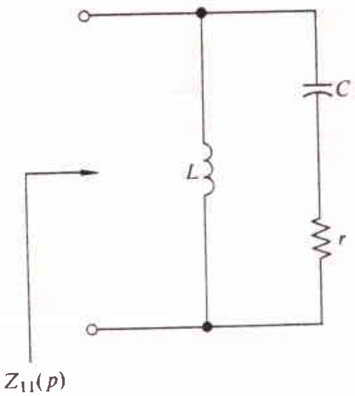


Fig. 2.3–7 Parallel resonant circuit with loss in series with the capacitor.

Figure 2.3–8 shows a pole-zero diagram of $Z_{11}(p)$ for the case where the poles form a complex conjugate pair. Again, and by reasoning similar to that employed when the loss is in a series with the inductor, it is readily shown that for $Q_C > 10$, where $Q_C = \omega_0/2\alpha = \omega_0 L/r = 1/\omega_0 rC$, the series loss r may be replaced by an equivalent parallel loss R_{eq} of value $Q_C^2 r$ with negligible effect on $Z_{11}(p)$.

In addition, it may be shown that if a loss r_1 appears in series with the inductor and a loss r_2 appears in series with the capacitor of a parallel resonant circuit, and if

$$Q_L = \frac{\omega_0 L}{r_1} > 10 \qquad \text{and} \qquad Q_C = \frac{\omega_0 L}{r_2} > 10,$$

where $\omega_0 = 1/\sqrt{LC}$, then the total loss may be represented as the parallel combination of $Q_L^2 r_1$, $Q_C^2 r_2$, and any existing shunt resistance. We shall explore this idea further in the following example.

Example 2.3–2 For the circuit shown in Fig. 2.3–9, determine the resonant radian frequency ω_0, the bandwidth, Q_T, and $Z_{11}(j\omega_0)$.

Solution. We observe first that $\omega_0 = 1/\sqrt{LC} = 10^7$ rad/sec; hence

$$Q_C = \frac{\omega_0 L}{r_2} = Q_L = \frac{\omega_0 L}{r_1} = 40 > 10.$$

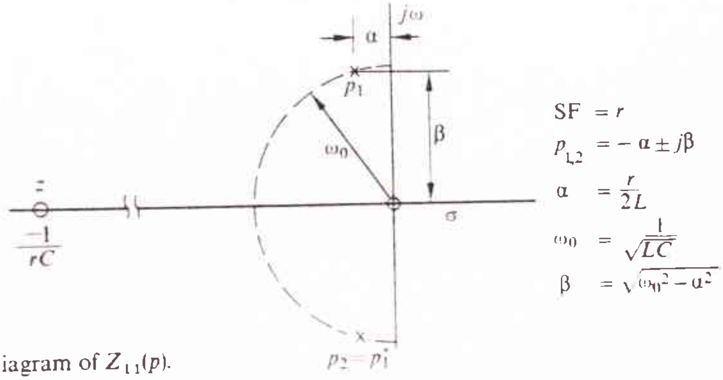


Fig. 2.3–8 Pole-zero diagram of $Z_{11}(p)$.

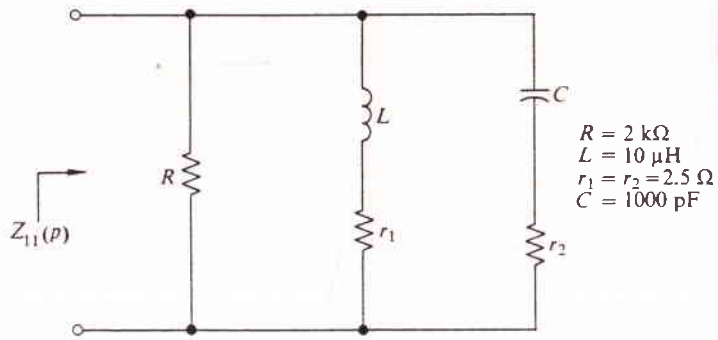


Figure 2.3–9

Consequently r_1 and r_2 may be replaced by equivalent parallel resistances each equal to $Q_C^2 r_1 = 4\ \text{k}\Omega$. The total shunt resistance is therefore the parallel combination of $2\ \text{k}\Omega$, $4\ \text{k}\Omega$, and $4\ \text{k}\Omega$, which of course is $1\ \text{k}\Omega = R_T$. Since for the equivalent parallel resonant circuit, L and C combine to produce an infinite impedance at ω_0, $Z_{11}(j\omega_0) = R_T = 1\ \text{k}\Omega$. Also $Q_T = R_T/\omega_0 L = 10$ and $\text{BW} = \omega_0/Q_T = 10^6$ rad/sec.

2.4 PARALLEL RESONANT "TRANSFORMERLIKE" NETWORKS

In this section we shall consider the parallel resonant circuit in which the loss appears across only a portion of the inductor or capacitor. Again we shall show that when Q is high (or, equivalently, the complex poles of the circuit are close to the imaginary axis), the circuit input impedance has essentially the same form as that of the parallel RLC circuit. The loss may therefore be modeled as an equivalent parallel resistor for which the results of Section 2.2 are directly applicable. In addition, we shall show that the tapped and loaded inductor or capacitor possesses many of the properties of an ideal transformer. Specifically, if the loading is light, the load across the tapped energy-storage element may be modeled by the identical load placed across the secondary of an ideal transformer whose primary is placed in parallel with the total energy-storage element. This representation is valid not only for the evaluation of the input impedance but also for the evaluation of the transfer impedance.

We begin our analysis with the circuit of Fig. 2.4–1, in which the loss in the form of G appears across a portion of the capacitor. We shall then generalize to the case where the loss appears across a portion of the inductor, and also (in Section 2.5) to the case where the loss appears across the secondary of a physical transformer with capacitively tuned primary.

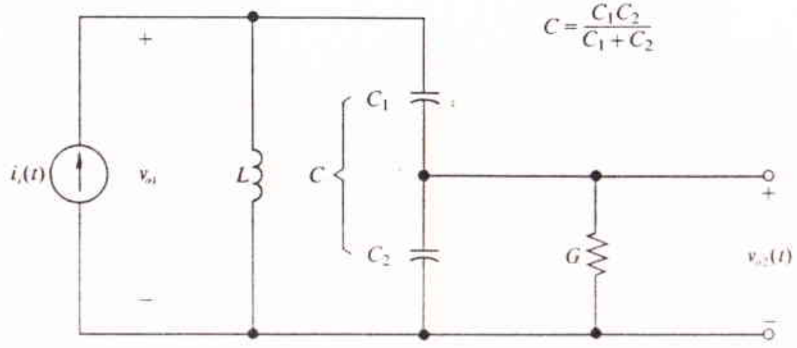


Fig. 2.4-1 Parallel resonant circuit with partially loaded capacitor.

For the circuit of Fig. 2.4–1 the input impedance is given by

$$Z_{11}(p) = \frac{V_{o1}(p)}{I_i(p)} = \frac{p\dfrac{1}{C}\left(p + \dfrac{G}{C_1 + C_2}\right)}{p^3 + p^2\dfrac{G}{C_2} + p\dfrac{1}{LC} + \dfrac{G}{LC_1C_2}}, \tag{2.4–1}$$

where $C = C_1C_2/(C_1 + C_2)$ is the series combination of C_1 and C_2. If the loading on C_2 is light, $Z_{11}(p)$ must have a pair of complex poles in addition to a real-axis pole, and thus may be written in the form

$$Z_{11}(p) = \frac{p\dfrac{1}{C}\left(p + \dfrac{G}{C_1 + C_2}\right)}{(p + \gamma)(p + \alpha - j\beta)(p + \alpha + j\beta)}, \tag{2.4–2}$$

where $p_1 = -\alpha + j\beta$ and $p_2 = -\alpha - j\beta$ are the complex conjugate poles and $p_3 = -\gamma$ is the real-axis pole. Figure 2.4–2 is a typical pole-zero diagram of $Z_{11}(p)$.

From Eq. (2.4–2) or from Fig. 2.4–2 it is apparent that the expression for $Z_{11}(p)$ reduces to the form of the input impedance of an equivalent parallel RLC circuit if $p_3 = z_2$ or, equivalently, if $\gamma = G/C_1 + C_2$. To determine under what circumstances

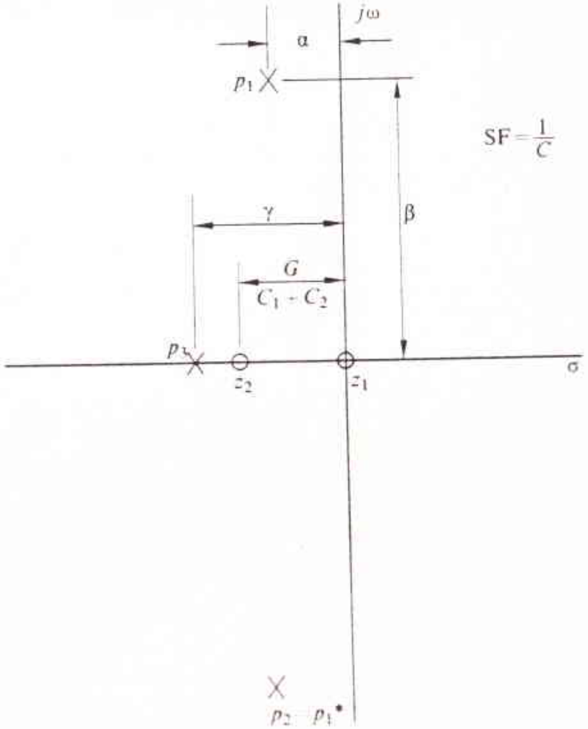


Fig. 2.4-2 Typical pole-zero diagram of $Z_{11}(p)$.

this pole-zero cancellation is achieved, and in addition the resultant values of α and β in terms of the circuit parameters of Fig. 2.4–1, we first write the denominator of Eq. (2.4–2) in the form

$$\begin{aligned}(p+\gamma)(p+\alpha-j\beta)(p+\alpha+j\beta)\\ = p^3+p^2(\gamma+2\alpha)+p(\alpha^2+\beta^2+2\alpha\gamma)+\gamma(\alpha^2+\beta^2).\end{aligned} \tag{2.4–3}$$

We then equate the corresponding powers of p in Eq. (2.4–3) and in the denominator of Eq. (2.4–1) to obtain the set of equations

$$\gamma + 2\alpha = G/C_2, \tag{2.4–4a}$$

$$\alpha^2+\beta^2+2\alpha\gamma = 1/LC \equiv \omega_0^2, \tag{2.4–4b}$$

$$\gamma(\alpha^2+\beta^2) = G/LC_1C_2. \tag{2.4–4c}$$

If, in addition, we define $\Omega = \omega_0^2/2\alpha\gamma$, then Eqs. (2.4–4a, b, and c) may be combined and rearranged in the form

$$2\alpha = \frac{n^2G}{C}\left(\frac{1-1/n\,\Omega}{1-1/\Omega}\right), \tag{2.4–5a}$$

$$\alpha^2+\beta^2 = \omega_0^2(1-1/\Omega), \tag{2.4–5b}$$

$$\gamma = \frac{G}{C_1+C_2}\left(\frac{1}{1-1/\Omega}\right), \tag{2.4–5c}$$

where $n = C_1/(C_1+C_2)$.

It now becomes clear from Eq. (2.4–5c) that if $\Omega > 100$ then the real-axis pole and zero of $Z_{11}(p)$ lie within 1 % of each other and, for all practical purposes, cancel to yield $Z_{11}(p)$ in the form of a parallel *RLC* circuit with $\alpha^2+\beta^2 \approx \omega_0^2$ and $2\alpha \approx (n^2G/C)(1-1/n\,\Omega)$.

With the aid of Eqs. (2.4–5a and c) we may also write

$$\Omega = \frac{\omega_0^2}{2\alpha\gamma} = \frac{Q_{T'}Q_E(1-1/\Omega)^2}{1-1/n\,\Omega}, \tag{2.4–6}$$

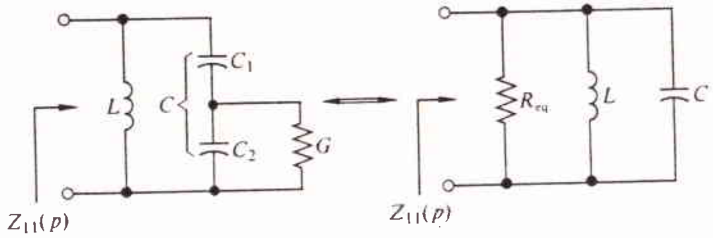
where $Q_{T'} = \omega_0 C/n^2G$ and $Q_E = \omega_0(C_1+C_2)/G$. For large values of Ω Eq. (2.4–6) reduces to

$$\Omega \approx Q_{T'}Q_E + 1/n. \tag{2.4–7}$$

Hence if $Q_{T'}Q_E$ (which may be directly determined in terms of the circuit parameters) is greater than 100, we can be sure that Ω is sufficiently large to effect the desired pole-zero cancellation. In addition, $Q_{T'}Q_E > 100$ ensures the accuracy (within 1 %) of the approximation of Eq. (2.4–7), from which we obtain

$$2\alpha = \frac{n^2G}{C}\left(1-\frac{1}{nQ_{T'}Q_E+1}\right) \tag{2.4–8}$$

$$Q_{T'}Q_E > 100, \quad Q_{T'} = \frac{\omega_0 C}{n^2 G}, \quad Q_E = \frac{\omega_0(C_1+C_2)}{G}$$



$$G_{eq} = Gn^2\left(1-\frac{1}{nQ_{T'}Q_E+1}\right), \quad n = \frac{C_1}{C_1+C_2}$$

Fig. 2.4–3 Parallel circuits with equivalent input impedances.

and in turn

$$Z_{11}(p) \approx \frac{p\dfrac{1}{C}}{p^2+2\alpha p+\alpha^2+\beta^2} \approx \frac{p\dfrac{1}{C}}{p^2+\dfrac{G_{eq}p}{C}+\dfrac{1}{LC}}, \tag{2.4–9}$$

where

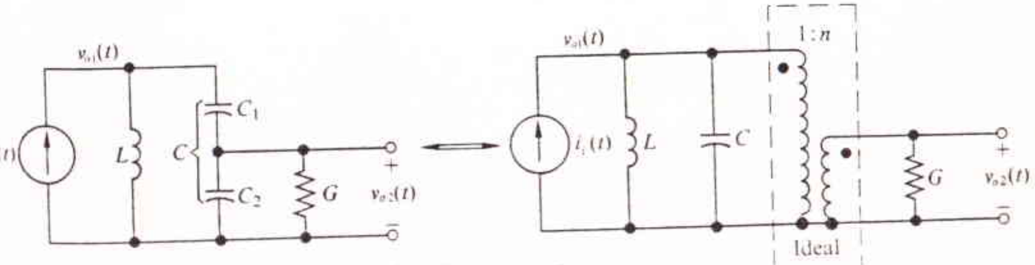
$$G_{eq} = n^2G\left(1-\frac{1}{nQ_{T'}Q_E+1}\right).$$

Equation (2.4–9) represents the input impedance of the equivalent *RLC* circuit shown in Fig. 2.4–3.

If $Q_{T'}Q_E > 100$ and, in addition, $nQ_{T'}Q_E$ is large, then G_{eq} reduces to the simplified form

$$G_{eq} = n^2G, \qquad nQ_{T'}Q_E \gg 1. \tag{2.4–10}$$

In particular, if $nQ_{T'}Q_E \geq 20$, then $G_{eq} = n^2G$ within 5 %, whereas if $nQ_{T'}Q_E > 100$, then $G_{eq} = n^2G$ within 1 %. (Note that since $n < 1$, $nQ_{T'}Q_E > 100$ ensures $Q_{T'}Q_E > 100$.) With this additional condition satisfied, we may transform the equivalent circuit shown in Fig. 2.4–3 into an alternative and more useful form, shown in Fig. 2.4–4. Clearly, if G is reflected through the ideal transform to obtain



$$Q_{T'} = \frac{\omega_0 C}{n^2G}, \quad Q_E = \frac{\omega_0(C_1+C_2)}{G}, \quad n = \frac{C_1}{C_1+C_2}$$

Fig. 2.4–4 Transformer model for resonant circuit with tapped and loaded capacitor.

$G_{eq} = n^2 G$, then the model of Fig. 2.4–3 results. However, the model of Fig. 2.4–4 is more versatile than the model of Fig. 2.4–3, since it provides a valid approximation for not only the input impedance $Z_{11}(p)$ but also the transfer impedance $Z_{12}(p) = V_{o2}(p)/I_i(p)$ for the case where $Q_E > 10$ and also $nQ_{T'} Q_E > 100$.

To demonstrate this property, we note that, in general,

$$Z_{12}(p) = \frac{V_{o2}(p)}{I_i(p)} = \frac{V_{o2}(p)}{V_{o1}(p)} \frac{V_{o1}(p)}{I_i(p)} = Z_{11}(p) H_v(p), \tag{2.4–11}$$

where

$$H_v(p) = \frac{V_{o2}(p)}{V_{o1}(p)} = \frac{C_1}{C_1 + C_2} \frac{p}{p + G/(C_1 + C_2)}$$

for the circuit of Fig. 2.4–1. A pole-zero diagram for $Z_{12}(p)$ is shown in Fig. 2.4–5. From this diagram and from arguments similar to those given in Section 2.3, it is clear that if $\omega_0/\gamma \gg 1$ and, in addition, $\omega_0/\alpha \gg 1$, then the zero at the origin effectively cancels the pole at $-\gamma$ when one evaluates the frequency or time response of $Z_{12}(p)$

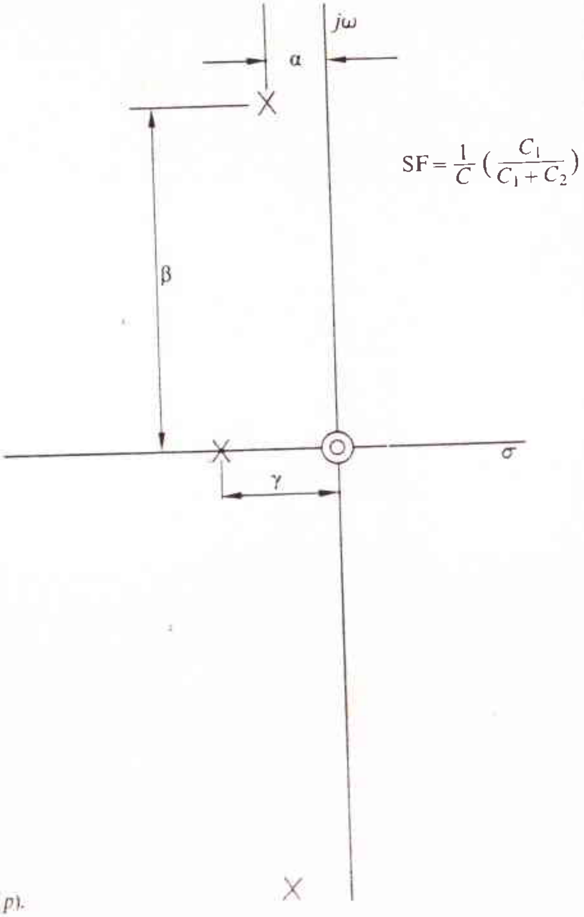


Fig. 2.4–5 Pole-zero pattern of $Z_{12}(p)$.

graphically; hence

$$Z_{12}(p) \approx \frac{C_1}{C_1 + C_2} Z_{11}(p) = n Z_{11}(p),$$

which is exactly the transfer impedance of the model of Fig. 2.4–4. Clearly the condition $nQ_{T'} Q_E > 100$ (which is required for the input impedance of the model of Fig. 2.4–4) results in $\omega_0/\alpha \approx Q_{T'}$ and $\omega_0/\gamma \approx Q_E$; hence the combined conditions $nQ_{T'} Q_E > 100$ and $Q_E > 10$ permit the use of the model of Fig. 2.4–4 for obtaining both $Z_{11}(p)$ and $Z_{12}(p)$ for the circuit of Fig. 2.4–1. Specifically, if $Q_E > 10$ (and $nQ_{T'} Q_E > 100$), then $|Z_{12}(j\omega)|$ obtained from the model is accurate within 1% of its actual peak value and the phase of $Z_{12}(j\omega)$ is accurate within 6° over the passband.

If the phase angle of $Z_{12}(j\omega)$ is critical or if $Q_E < 10$ while $Q_{T'} Q_E > 100$, then $Z_{11}(j\omega)$ should be obtained from the model of Fig. 2.4–3 and $Z_{12}(j\omega)$ should be obtained by multiplying $Z_{11}(j\omega)$ by $H_v(j\omega)$, which is determined exactly from Eq. (2.4–11).

At this point it is worthwhile to interpret physically the parameters n, C, $Q_{T'}$, and Q_E employed in the models of Figs. 2.4–3 and 2.4–4. Clearly n is the voltage division ratio of the two series capacitors with the load across C_2 removed ($G = 0$); that is,

$$n = \left.\frac{v_{o2}(t)}{v_{o1}(t)}\right|_{G=0}.$$

The capacitor C is just the total capacitance shunting L in the circuit of Fig. 2.4–1 obtained with $G = 0$. In addition, $Q_{T'} = \omega_0 C/n^2 G$ is the Q of the model of Fig. 2.4–4, whereas $Q_E = \omega_0(C_1 + C_2)/G$ is the ratio of the output resistance $(1/G)$ to the reactance at ω_0 which shunts G, evaluated with the input voltage $v_o(t)$ of the circuit of Fig. 4.2–1 reduced to zero (i.e., with the input shorted). We shall now use the same physical interpretation to obtain the corresponding parameters for a parallel RLC circuit with a tapped and loaded inductor.

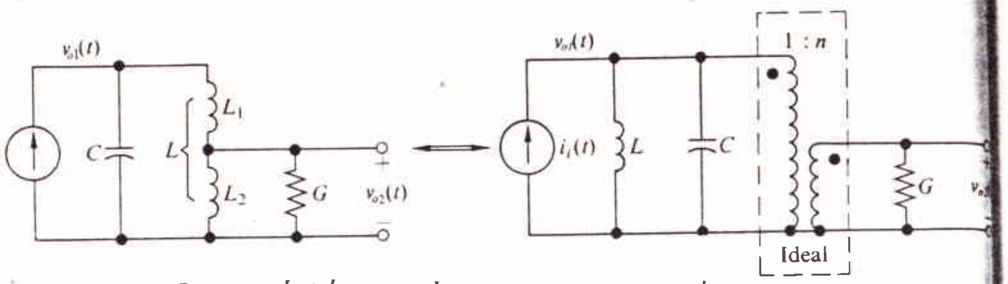
Consider the parallel RLC circuit of Fig. 2.4–6. By a procedure similar to that employed with the circuit of Fig. 2.4–1, we can show that the transformer model, also shown in Fig. 2.4–6, may be employed to obtain expressions for $Z_{11}(p)$ and $Z_{12}(p)$ provided that

$$nQ_{T'} Q_E > 100 \quad \text{and} \quad Q_E > 10,$$

where in this case $Q_{T'} = \omega_0 C/n^2 G$, $Q_E = \omega_0 L_1 L_2/G(L_1 + L_2)$, $\omega_0 = 1/\sqrt{LC}$, $L = L_1 + L_2$, and $n = L_2/(L_1 + L_2)$. Here again we observe that

$$n = \left.\frac{v_{o2}(t)}{v_{o1}(t)}\right|_{G=0},$$

that L is the inductance shunting C with $G = 0$, that $Q_{T'}$ is the Q of the model of Fig. 2.4–6, and that Q_E is the ratio of $1/G$ to the reactance at ω_0 shunting G, evaluated with $v_{o1}(t) = 0$. This set of conditions is valid for any of the transformerlike networks shown in Table 2.5–1 and thus provides a handy mnemonic rule. If for the circuit of Fig. 2.4–6 $Q_{T'} Q_E > 100$ but $Q_E < 10$, $Z_{11}(p)$ may be evaluated from an equivalent



$$Q_T = \frac{\omega_0 C}{n^2 G},\quad Q_E = \frac{L_1 + L_2}{\omega_0 G L_1 L_2},\quad n = \frac{L_2}{L_1 + L_2},\quad L = L_1 + L_2,\quad \omega_0 = \frac{1}{\sqrt{LC}}$$

Fig. 2.4–6 Parallel *RLC* circuit with tapped and loaded inductor and its transformer model.

parallel *RLC* circuit with a shunt resistance $R_{\text{eq}} = 1/G_{\text{eq}}$, where again

$$G_{\text{eq}} = n^2 G\left(1 - \frac{1}{nQ_T \cdot Q_E + 1}\right). \tag{2.4–12}$$

In addition, $Z_{12}(p)$ may be evaluated as $Z_{12}(p) = Z_{11}(p)H_v(p)$, where

$$H_v(p) = \frac{1}{L_1 G}\,\frac{1}{p + (L_1 + L_2)/L_1 L_2 G}. \tag{2.4–13}$$

Shunt Input Resistance

If an additional shunt input resistance R_L is placed across the parallel *RLC* circuit with a tapped and loaded capacitor, as shown in Fig. 2.4–7, it is apparent that the input impedance

$$Z'_{11}(p) = \frac{V_{o1}(p)}{I_i(p)} = R_L \| Z_{11}(p) \quad \text{and} \quad Z'_{12}(p) = \frac{V_{o2}(p)}{V_{o1}(p)} = Z'_{11}(p)H_v(p),$$

where $Z_{11}(p)$ and $H_v(p)$ are the input impedance and transfer voltage function of the circuit of Fig. 2.4–1. Consequently, if $nQ_T \cdot Q_E > 100$, $Z_{11}(p)$ may be obtained from

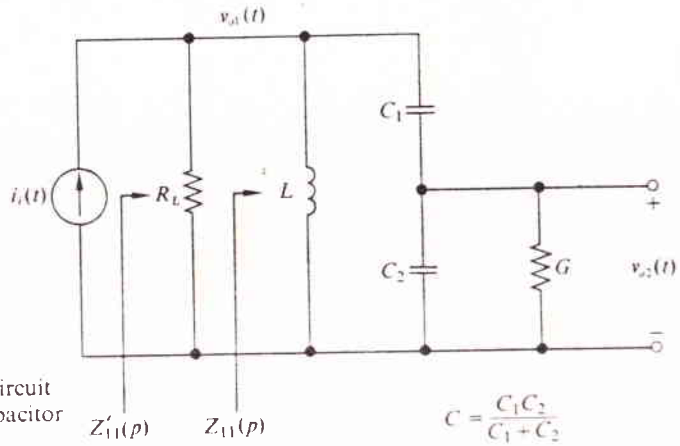
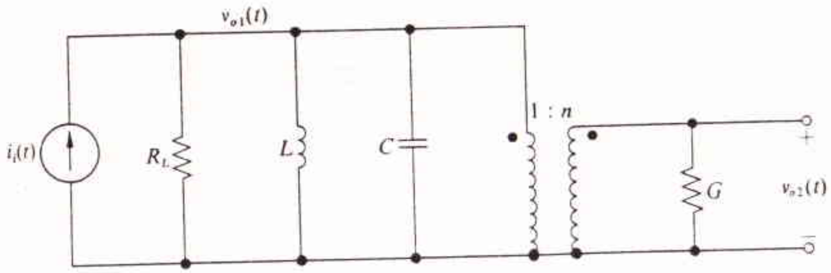


Fig. 2.4–7 Parallel *RLC* circuit with tapped and loaded capacitor and an input resistance.

$$C = \frac{C_1 C_2}{C_1 + C_2}$$



$$Q_T = \frac{\omega_0 C}{G_L + n^2 G} > 10,\quad nQ_T \cdot Q_E > 100,\quad Q_E > 10,\quad n = \frac{C_1}{C_1 + C_2}$$

Fig. 2.4–8 Model for circuit of Fig. 2.4–7 used to obtain $Z'_{11}(p)$ and $Z'_{12}(p)$.

an equivalent parallel *RLC* circuit with a shunt resistance of $1/n^2 G$, and in turn $Z'_{11}(p)$ may be obtained from the equivalent parallel *RLC* circuit with shunt resistance $1/(n^2 G + G_L)$, shunt capacitance C, and shunt inductance L. Clearly, for this resultant parallel *RLC* circuit

$$Q_T = \frac{\omega_0 C}{n^2 G + G_L} = \frac{\omega_0}{2\alpha'}, \tag{2.4–14}$$

where $-\alpha'$ is the real part of the poles of $Z'_{11}(p)$.

If both $\omega_0/2\alpha' = Q_T > 10$ and $\omega_0/\gamma > 10$, where γ is the pole of $H_v(p)$, then again $Z'_{12}(p) \approx Z'_{11}(p)n$. If $nQ_T \cdot Q_E > 100$, then $\omega_0/\gamma \approx Q_E$; thus if $nQ_T/Q_E > 100$, $Q_E > 10$, and $Q_T > 10$, then the model shown in Fig. 2.4–8 may be used to obtain both $Z'_{11}(p)$ and $Z'_{12}(p)$ for the circuit of Fig. 2.4–7. The same arguments apply to the parallel *RLC* circuit with a tapped and loaded inductor.

Example 2.4–1 For the circuit shown in Fig. 2.4–9, determine an expression for $v_{o1}(t)$ and $v_{o2}(t)$ as well as the circuit bandwidth.

Solution. If we assume $nQ_T \cdot Q_E > 100$ and $Q_E > 10$, we may replace the original circuit by the model shown in Fig. 2.4–10, for which

$$n = \frac{C_1}{C_1 + C_2} = \frac{1}{2}, \qquad C = \frac{C_1 C_2}{C_1 + C_2} = 1000\ \text{pF},$$

$$\omega_0 = \frac{1}{\sqrt{LC}} = 10^7\ \text{rad/sec}, \qquad Z_{11}(j\omega_0) = R_{\text{eq}} = \frac{500\ \Omega}{n^2} = 2\ \text{k}\Omega,$$

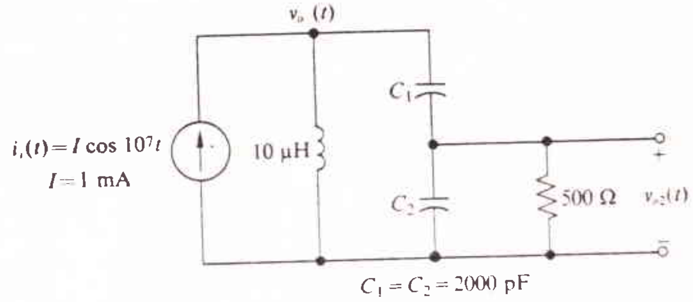


Figure 2.4–9

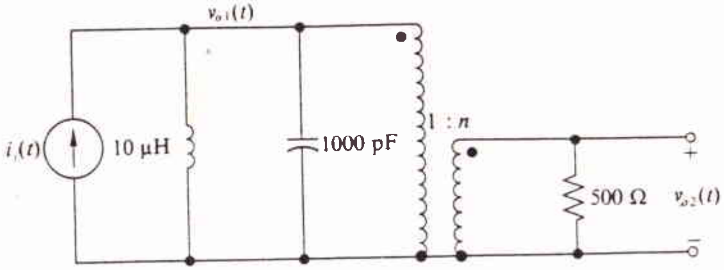


Figure 2.4–10

and

$$Z_{12}(j\omega_0) = nZ_{11}(j\omega_0) = 1\ \mathrm{k\Omega}.$$

In addition we note that

$$Q_{T'} = \frac{R_{\mathrm{eq}}}{\omega_0 L} = 20, \qquad Q_E = (500\ \Omega)\omega_0(C_1 + C_2) = 20, \qquad \text{and} \qquad nQ_{T'}Q_E = 200;$$

these values justify the use of the model.

Since the current drive is at the resonant frequency ω_0, where the input impedance appears purely resistive $[Z_{11}(j\omega_0) = R_{\mathrm{eq}}]$, $v_{o1}(t)$ may be written directly in the form

$$v_{o1}(t) = IR_{\mathrm{eq}} \cos 10^7 t = 2\ \mathrm{V} \cos 10^7 t,$$

and hence

$$v_{o2}(t) = nv_{o1}(t) = 1\ \mathrm{V} \cos 10^7 t.$$

Finally, since the only loading in the circuit is across C_2, $Q_{T'} = Q_T$ and BW $= \omega_0/Q_T = 5 \times 10^5$ rad/sec.

Example 2.4–2 For the circuit shown in Fig. 2.4–11 determine values for C_1, C_2, and L such that the circuit resonates at $f_0 = 16$ MHz ($\omega_0 = 10^8$ rad/sec) with a bandwidth of 1.06 MHz [BW = (10^7/1.5) rad/sec] and achieves maximum signal transmission to R_L at resonance.

Solution. If we again assume $nQ_{T'}Q_E > 100$, $Q_E > 10$, and $Q_T > 10$, we may replace the original circuit by the model shown in Fig. 2.4–12. Since at resonance the inductor and capacitor combine to produce an open circuit, maximum signal (or power) is transmitted to R_L in the purely resistive network when n is chosen to "match" R_L to R_s, i.e., when

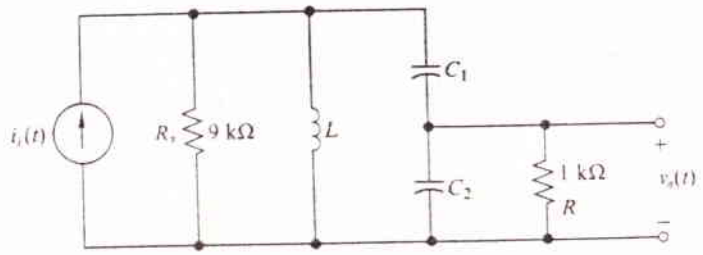
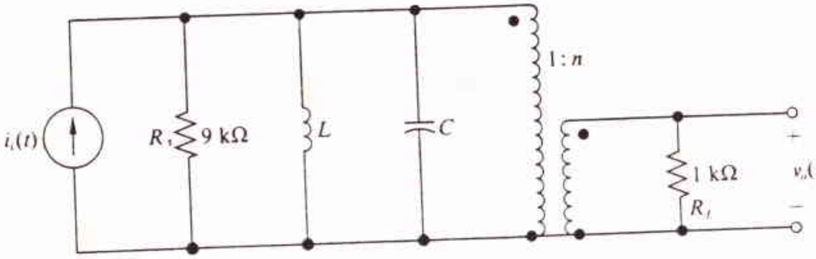


Figure 2.4–11



$$n = \frac{C_1}{C_1 + C_2}, \quad C = \frac{C_1 C_2}{C_1 + C_2}, \quad \omega_0 = \frac{1}{\sqrt{LC}}$$

Figure 2.4–12

$$n = \sqrt{R_L/R_s} = \sqrt{1\ \mathrm{k\Omega}/9\ \mathrm{k\Omega}} = 1/3 = C_1/(C_1 + C_2).$$

With this value of n, R_L is reflected through the ideal transformer as a 9 kΩ resistor; thus the total input resistance at resonance is 4.5 kΩ. Recalling that BW $= 1/RC$ for a parallel RLC circuit, we obtain

$$C = \frac{C_1 C_2}{C_1 + C_2} = \frac{1}{(4.5\ \mathrm{k\Omega})\mathrm{BW}} = \frac{100}{3}\mathrm{pF}.$$

Replacing $C_1/(C_1 + C_2)$ by $n = \frac{1}{3}$, we obtain

$$C_2 = \frac{C}{n} = 100\ \mathrm{pF}, \qquad C_1 = \left(\frac{1}{n} - 1\right)C_2 = 50\ \mathrm{pF},$$

and finally

$$L = \frac{1}{\omega_0^2 C} = 30\ \mu\mathrm{H}.$$

At this point the original assumption must be checked. It is apparent that

$$Q_{T'} = \frac{1\ \mathrm{k\Omega}}{n^2}\omega_0 C = 30 \qquad \text{and} \qquad Q_E = (1\ \mathrm{k\Omega})\omega_0(C_1 + C_2) = 15;$$

hence $nQ_{T'}Q_E = 150$. In addition, we note that for the circuit in this example $Q_T = \omega_0/\mathrm{BW} = 15$; thus the use of the model is justified.

Example 2.4–3 For the circuit shown in Fig. 2.4–13, determine $Z_{11}(j\omega_0)$ and $Z_{12}(j\omega_0)$.

Solution. For the circuit shown, $\omega_0 = 1/\sqrt{LC} = 10^7$ rad/sec and $n = L_2/(L_1 + L_2) = 1/100$. Hence

$$Q_{T'} = \frac{1\ \Omega}{n^2 \omega_0 L} = 100, \qquad Q_E = \frac{1\ \Omega}{\omega_0 \dfrac{L_1 L_2}{L_1 + L_2}} \approx \frac{1\ \Omega}{\omega_0 L_2} = 1, \qquad \text{and} \qquad Q_T Q_E = 100.$$

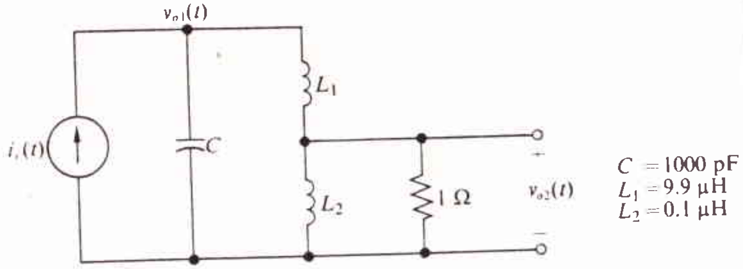


Figure 2.4–13

Since $Q_T \cdot Q_E = 100$ but $nQ_T \cdot Q_E = 1$,

$$Z_{11}(j\omega_0) = R_{\text{eq}} = \frac{1}{n^2 G\left(1 - \dfrac{1}{nQ_T \cdot Q_E + 1}\right)} = \frac{1\ \Omega}{\dfrac{n^2}{2}} = 20\ \text{k}\Omega$$

(cf. Eq. 2.4–12). On the other hand, since $Q_E < 10$, the transformer model shown in Fig. 2.4–6 may not be used to determine $Z_{12}(p)$. However, the exact expression for $Z_{12}(p)$ is given by

$$Z_{12}(p) = Z_{11}(p)H_v(p) = Z_{11}(p)\frac{\dfrac{1}{L_1 G}}{p + \dfrac{L_1 + L_2}{L_1 L_2 G}}. \tag{2.4–15}$$

If we substitute $j\omega_0 = j10^7$ rad/sec for p in Eq. (2.4–16), we obtain $Z_{12}(j\omega_0)$ in the form

$$Z_{12}(j\omega_0) = (20\ \text{k}\Omega)\frac{1}{100(1 + j)},$$

from which we obtain

$$|Z_{12}(j\omega_0)| = \frac{200}{\sqrt{2}}\ \Omega \qquad \text{and} \qquad \arg Z_{12}(j\omega_0) = -45°.$$

2.5 PARALLEL RESONANT TRANSFORMERS

In applications where a narrowband circuit is needed to isolate the load from the driving source or to provide a 180° phase inversion between the input signal and the load at resonance, a physical transformer must be employed in the parallel resonant circuit as shown in Fig. 2.5–1. In addition, the physical transformer must be employed if a "step-up" turns ratio is required between the driving source and the load at resonance for the purpose of matching a high load impedance to a small source impedance. The "transformerlike" networks discussed in the previous section are capable of providing only a "step-down" turns ratio.

The physical transformer may also be employed in a parallel resonant circuit as an auto transformer, shown in Fig. 2.5–2. In this configuration the transformer does not provide isolation, 180° phase inversion, or a "step-up" turns ratio at resonance. Its usefulness stems from the fact that it is a more realistic model for the inductive "transformerlike" parallel resonant circuit, since it includes the effect of magnetic coupling between the two inductors in the form of mutual inductance. The auto-transformer representation is especially useful when the two inductors are wound on the same core or obtained by providing a tap on a single inductor. As is shown in Fig. 2.5–2, L_1 and L_2 are defined differently in this case than when the two inductors are independent. This is done so that the networks of Figs. 2.5–1 and 2.5–2 can be analyzed by means of the same model.

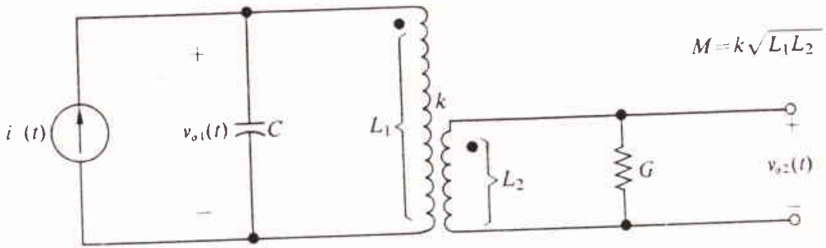


Fig. 2.5–1 Transformer with tuned primary and loaded secondary.

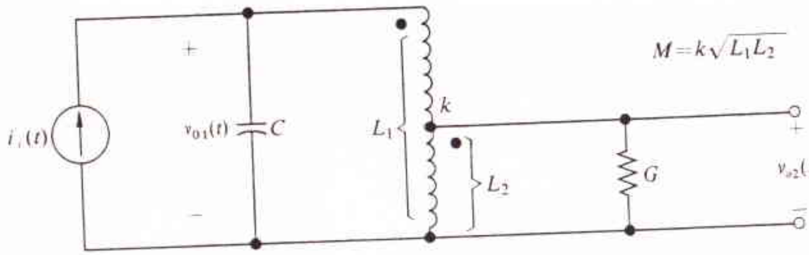


Fig. 2.5–2 Auto transformer with tuned primary and loaded secondary.

To obtain expressions for the input and transfer impedances of the two parallel resonant transformer networks, the results of Section 2.4 may be applied directly. This becomes apparent if we replace the physical transformers of Figs. 2.5–1 and 2.5–2 by their common "two-inductor" terminal equivalent model (which is derived in the appendix to this chapter) as shown in Fig. 2.5–3. Clearly the ideal transformer reflects a net conductance of $G/a^2 = GL_2/k^2L_1$ across the inductor k^2L_1 and thus the

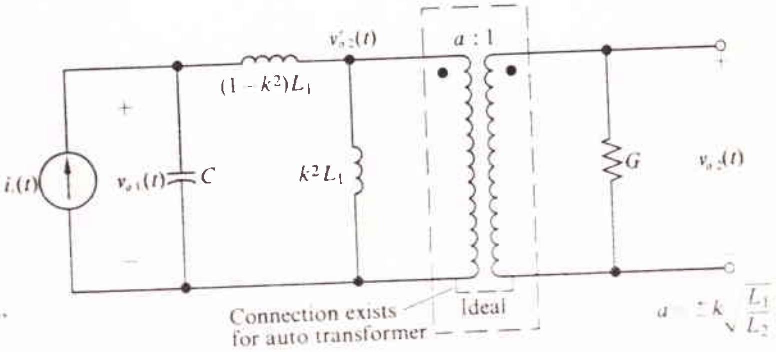


Fig. 2.5–3 Circuits of Figs. 2.5–1 and 2.5–2 with transformers replaced by their "two-inductor" model.

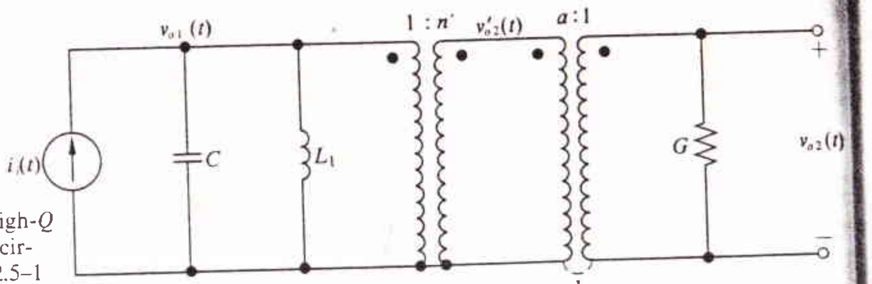


Fig. 2.5–4 High-Q model for the circuits of Figs. 2.5–1 and 2.5–2.

$$n'=k^2,\quad a=k\sqrt{\frac{L_1}{L_2}},\quad Q_{T'}=\frac{\omega_0 C}{\left(\frac{M}{L_1}\right)G},\quad Q_E=\frac{1}{\omega_0 L_2(1-k^2)G}$$

circuit becomes identical to the circuit of Fig. 2.4–6. Consequently, if

$$n'Q_{T'}Q_E > 100\dagger \qquad \text{and} \qquad Q_E > 10,$$

where

$$n' = k^2, \qquad Q_{T'} = \frac{\omega_0 C}{k^2 G/a^2} = \frac{\omega_0 C}{(M/L_1)^2 G}, \qquad \text{and} \qquad Q_E = \frac{1}{\omega_0 L_2(1-k^2)G},$$

then we may replace the circuit to the left of the ideal transformer by its equivalent transformer model as shown in Fig. 2.5–4 (cf. Fig. 2.4–6).

If we now combine the two ideal transformers in cascade to obtain a single ideal transformer with a transformation ratio $n = n'/a = k\sqrt{L_2/L_1} = M/L_1$, the model of Fig. 2.5–4 reduces to the desired model for determining $Z_{11}(p)$ and $Z_{12}(p)$, shown in Fig. 2.5–5. Here again we observe for the circuits of Figs. 2.5–1 and 2.5–2 that

$$n = \left.\frac{v_{o2}(t)}{v_{o1}(t)}\right|_{G=0},$$

that L_1 is the total inductance shunting C with $G = 0$, that Q_E is the ratio of the resistance $(1/G)$ to the total reactance at ω_0 shunting G with $v_{o1}(t) = 0$, and that $Q_{T'}$ is the Q of the model of Fig. 2.5–5.

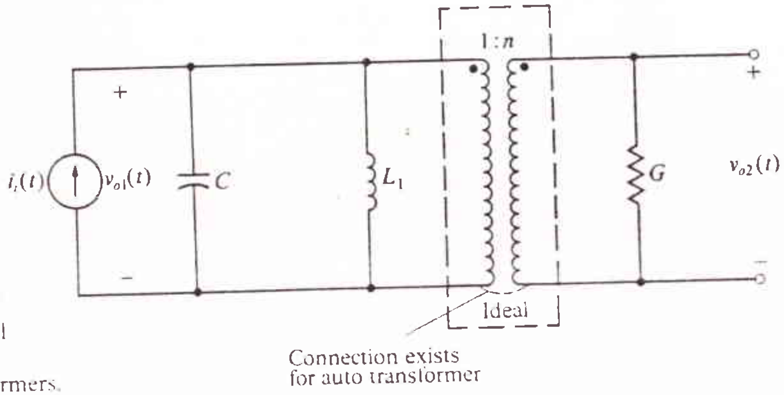


Fig. 2.5–5 Final model for parallel resonant circuits including transformers.

---

† We use n' in lieu of n because n is being reserved for the overall transformer ratio.

Table 2.5—1 Equivalent circuits and pertinent relationships for parallel resonant "transformer-like" networks

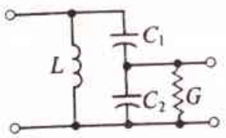
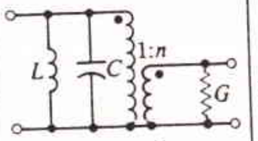
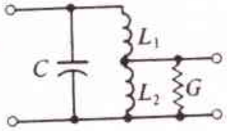
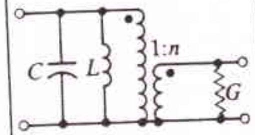
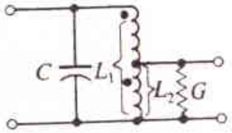
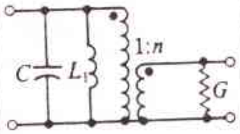
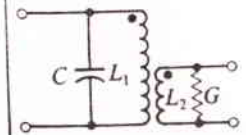
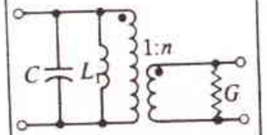
Circuit	Model for determining $Z_{11}(p)$ and $Z_{12}(p)$	n	ω_0	Q_T	Q_E	
	 $C = \frac{C_1 C_2}{C_1+C_2}$	$\frac{C_1}{C_1+C_2}$ $n'=n$	$\frac{1}{\sqrt{LC}}$	$\frac{\omega_0 C}{n^2 G}$	$\frac{\omega_0(C_1+C_2)}{G}$	Limitations for $Z_{11}(p)$: $n'Q_T{}'Q_E > 100$ Limitations for $Z_{12}(p)$: $n'Q_T{}'Q_E > 100$ and $Q_E > 10$
	 $L = L_1 + L_2$	$\frac{L_2}{L_1+L_2}$ $n'=n$	$\frac{1}{\sqrt{LC}}$	$\frac{\omega_0 C}{n^2 G}$	$\frac{L_1+L_2}{\omega_0 L_1 L_2 G}$	
		$\frac{M}{L_1}$ $n'=k^2$	$\frac{1}{\sqrt{L_1 C}}$	$\frac{\omega_0 C}{n^2 G}$	$\frac{1}{\omega_0 L_2(1-k^2)G}$	
		$\frac{M}{L_1}$ $n'=k^2$	$\frac{1}{\sqrt{L_1 C}}$	$\frac{\omega_0 C}{n^2 G}$	$\frac{1}{\omega_0 L_2(1-k^2)G}$	

Table 2.5–1 summarizes these properties that we have discussed.

Since the turns ratio n of an auto transformer in a high-Q parallel resonant circuit may be expressed in terms of an unloaded voltage ratio, the physical construction of the auto transformer is quite straightforward. One selects (or winds) an inductor whose inductance is the desired value of L_1, places a voltage source $V_1 \cos \omega_0 t$ across the entire inductor, and moves a probe along the length of the inductor. At the point where the voltage $V_2 \cos \omega_0 t$ measured between the probe and the bottom of the inductor is equal to nV_1, a lead is soldered and the construction

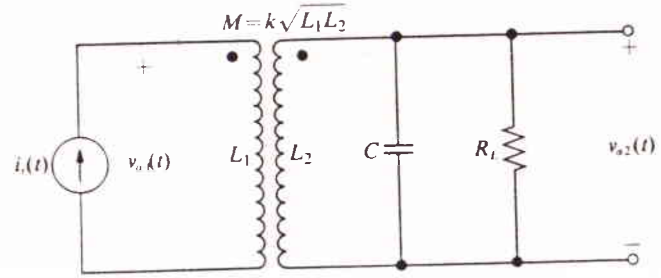


Fig. 2.5-6 Parallel resonant transformer with tuned secondary.

of the auto transformer is complete. Clearly this transformer resonates at $\omega_0 = 1/\sqrt{L_1C}$ and possesses a turns ratio of n.

Before ending our discussion of parallel resonant transformer circuits, we should note that no restrictions on the coefficient of coupling of the physical transformer were required to derive the high-Q model of Fig. 2.5-5. However, if k approaches unity, Q_E approaches infinity and the model of Fig. 2.5-5 becomes an exact model for the parallel resonant transformer, regardless of the value of G. If we had decided to restrict our attention to closely coupled transformers, we could have obtained the model of Fig. 2.5-5 directly by letting $k = 1$ in the model of Fig. 2.5-3 and noting that $n = 1/a = \pm\sqrt{L_2/L_1} = \pm M/L_1$.

Transformers with Tuned Secondaries

In many applications the secondary, rather than the primary, of a transformer is tuned, as shown in Fig. 2.5-6. Such coupling is usually employed to minimize the effect of any capacitance shunting $i_i(t)$ (for instance, collector capacitance) on the tuning of the resonant circuit. Specifically, if the transformer steps up impedance by a factor of a^2, the effective capacitance shunting C is $1/a^2$ times the capacitance shunting $i_i(t)$. If $a = 10$ and the current source capacity is 2 pF, only 0.02 pF reflects through to load C.

The tuned secondary is also employed with transistor IF amplifiers to keep the collector impedance and, in turn, the collector voltage small. This low impedance level not only reduces the "Miller" capacitance seen at the transistor input, but also reduces the effect of output tuning on the input impedance. Such isolation is essential if stages are to be tuned independently to obtain an overall IF transfer function.

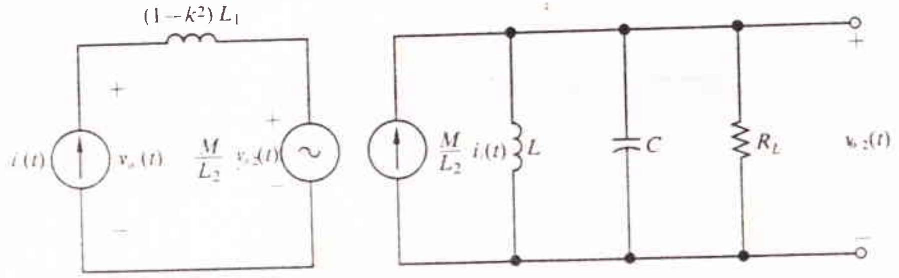


Fig. 2.5-7 Hybrid controlled-source model replacing transformer of Fig. 2.5-6

To obtain expressions for the input and transfer impedance of the circuit shown in Fig. 2.5-6, we first replace the transformer by its hybrid controlled-source model (derived in the appendix to this chapter), as shown in Fig. 2.5-7. With the model (which is not based on a high-Q approximation) in place, it is immediately apparent that $Z_{12}(p) = V_{o2}(p)/I_i(p)$ is the input impedance of a parallel RLC circuit multiplied by M/L_2; specifically,

$$Z_{12}(p) = \frac{\frac{M}{L_2}p\frac{1}{C}}{p^2 + \frac{p}{R_LC} + \frac{1}{L_2C}}. \tag{2.5-1}$$

In addition,

$$\begin{aligned} Z_{11}(p) &= p(1-k^2)L_1 + \frac{M}{L_2}Z_{12}(p) \\ &= p(1-k^2)L_1 + \left(\frac{M}{L_2}\right)^2 \frac{p\frac{1}{C}}{p^2 + \frac{p}{R_LC} + \frac{1}{L_2C}}. \end{aligned} \tag{2.5-2}$$

We therefore observe that the effective transformation ratio is M/L_2, the factor by which $i_i(t)$ is reduced when reflected into the secondary and the factor squared by which the parallel RLC circuit is reflected into the primary. We also observe that if $i_i(t)$ is a periodic input current of period $T = 2\pi/\omega_0$, then $v_{o2}(t)$ will be a sinusoid with a frequency ω_0 if $Q_T = \omega_0R_LC \gg 1$, since the tuned circuit will extract only the fundamental component of $i_i(t)$. On the other hand, because of the series inductance $(1-k^2)L_1$, $v_{o1}(t)$ will be rich in harmonic content.

Example 2.5-1 For the circuit shown in Fig. 2.5-8 derive an expression for $Z_{12}(p)$ and $Z_{13}(p)$ with the assumption that the circuit is a high-Q one.

Solution. Using the hybrid controlled-source model, we may first reflect $i_i(t)$ through the transformer to obtain a source of $(M/L_2)i_i(t)$ shunting R_L. If we now assume

$$nQ_T \cdot Q_E > 100, \qquad Q_E > 10, \qquad \text{and} \qquad Q_T > 10,$$

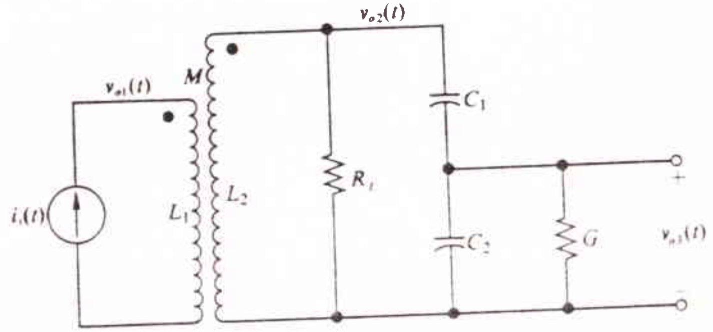


Figure 2.5-8

where

$$n = \frac{C_1}{C_1 + C_2}, \qquad Q_{T'} = \frac{\omega_0 C}{n^2 G}, \qquad Q_T = \frac{\omega_0 C}{n^2 G + G_L},$$

$$Q_E = \frac{\omega_0 (C_1 + C_2)}{G}, \qquad \text{and} \qquad C = \frac{C_1 C_2}{C_1 + C_2},$$

then we may replace the tapped capacitor by an ideal transformer to obtain the model shown in Fig. 2.5–9. From this circuit it is immediately apparent that

$$Z_{12}(p) = \frac{\dfrac{M}{L_2} p \dfrac{1}{C}}{p^2 + \dfrac{(n^2 G + G_L)p}{C} + \dfrac{1}{L_2 C}} \tag{2.5–3}$$

and

$$Z_{13}(p) = \frac{n \dfrac{M}{L_2} p \dfrac{1}{C}}{p^2 + \dfrac{(nG + G_L)p}{C} + \dfrac{1}{L_2 C}}.$$

Networks of the form shown in Fig. 2.5–9 are usually employed when a single tuned response is required, as well as a low impedance across $i_i(t)$ and an overall voltage step-down ratio. In addition, such networks are employed to keep stray capacitance across both $i_i(t)$ and G from affecting the tuning of the parallel RLC circuit.

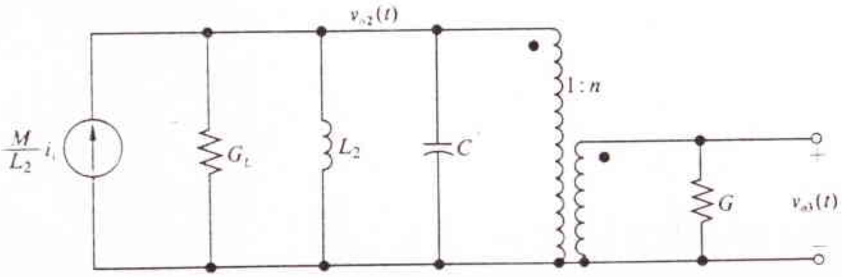


Figure 2.5–9

$$C = \frac{C_1 C_2}{C_1 + C_2}, \quad n = \frac{C_1}{C_1 + C_2}$$

2.6 THREE-WINDING PARALLEL RESONANT TRANSFORMER

In this section we shall develop a model for the three-winding parallel resonant transformer shown in Fig. 2.6–1. Since this circuit is a fundamental component not only in the antenna stage, but also in the oscillator and IF amplifier stages of superheterodyne receivers, a simplified model is required to facilitate its analysis and design. In addition, the model aids in the understanding of the physical operation of the transformer network.

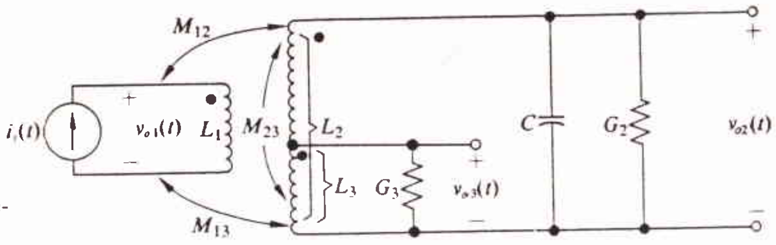


Fig. 2.6–1 Three-winding parallel resonant transformer.

We begin by presenting an equivalent circuit for the three-winding transformer. We then incorporate this model into the circuit of Fig. 2.6–1 and, with the results of the previous section and the assumption of high-Q, reduce the circuit to its final simplified model.

The most general model for the three-winding transformer is quite complicated. However, if we place the restriction on the transformer that

$$\frac{M_{12}}{L_2} = \frac{M_{13}}{M_{23}}, \tag{2.6–1}$$

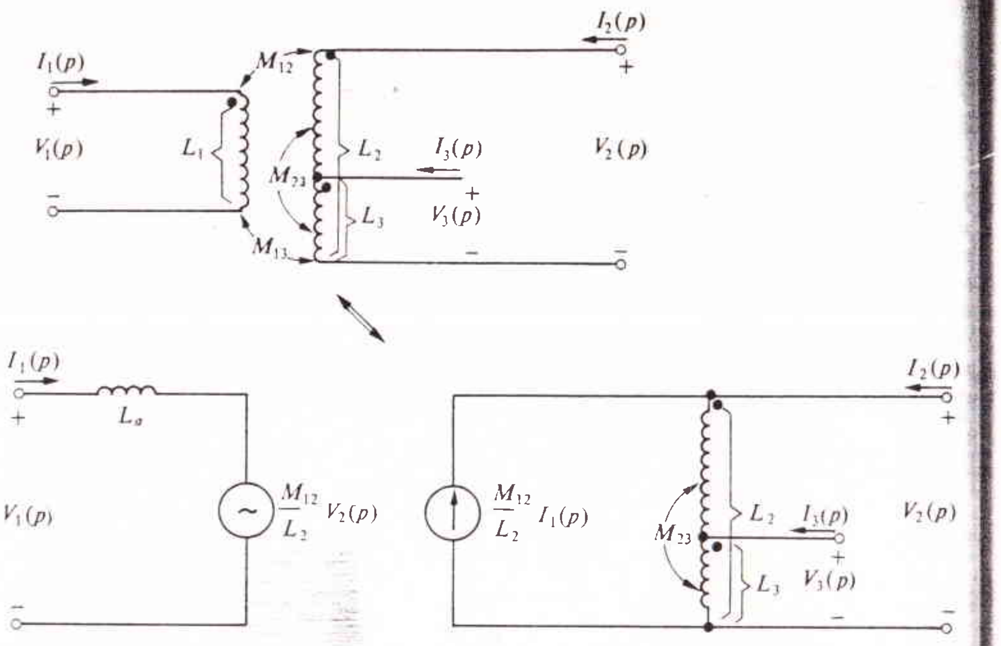
where $M_{ij} = M_{ji}$ is the mutual inductance between the ith and the jth windings, then a straightforward and useful model may be obtained. Equation (2.6–1) can be cast in the equivalent form

$$k_{12} k_{23} = k_{13} \tag{2.6–2}$$

by noting that $k_{ij} = M_{ij}/\sqrt{L_i L_j}$, where k_{ij} is the coefficient of coupling between the ith and jth windings. Although, in general, Eq. (2.6–1) is not always exactly satisfied, it is approximately satisfied in a large number of practical cases. In particular, if L_2 and L_3 are closely coupled (they are usually wound on the same core), then $k_{23} \approx 1$ and, of course, $k_{12} \approx k_{13}$; hence Eq. (2.6–1) is valid. In addition, if the transformer is placed in a high-Q circuit, it can be shown, by the techniques developed in Section 2.4, that even if Eq. (2.6–1) is not satisfied, the impedances obtained on the assumption that it is are accurate within a few percent. The proof of this statement is left to the interested reader. With the restriction of Eq. (2.6–1), the three-winding transformer has the terminal equivalent model shown in Fig. 2.6–2. This equivalence is readily demonstrated by showing that both the transformer and its model have the identical set of defining equations:

$$\begin{aligned} V_1 &= pL_1 I_1 + pM_{12} I_2 + pM_{13} I_3, \\ V_2 &= pM_{12} I_1 + pL_2 I_2 + pM_{23} I_3, \\ V_3 &= pM_{13} I_1 + pM_{23} I_2 + pL_3 I_3. \end{aligned} \tag{2.6–3}$$

If the transformer is now replaced by its model in the circuit of Fig. 2.6–1, the equivalent circuit shown in Fig. 2.6–3 results. It is apparent that the right-hand portion of the circuit in Fig. 2.6–3 is simply the parallel resonant auto transformer



$$L_a = L_1\left(1 - \frac{M_{12}\,M_{13}}{L_1\;M_{23}}\right) = L_1\left(1 - \frac{k_{12}\;k_{13}}{k_{23}}\right)$$

Fig. 2.6–2 Terminal equivalent circuit for three-winding transformers with $M_{12}/L_2 = M_{13}/M_{23}$.

considered in Section 2.5: hence we may apply the results of that section directly. Specifically, if

$$n'Q_T Q_E > 100 \quad (n' = k_{23}), \qquad Q_E > 10, \qquad \text{and} \qquad Q_T > 10,$$

where

$$Q_{T'} = \frac{\omega_0 C}{G_3}\left(\frac{L_2}{M_{24}}\right)^2, \qquad Q_E = \frac{1}{\omega_0 L_3(1 - k_{23}^2)G_3},^{\dagger}$$

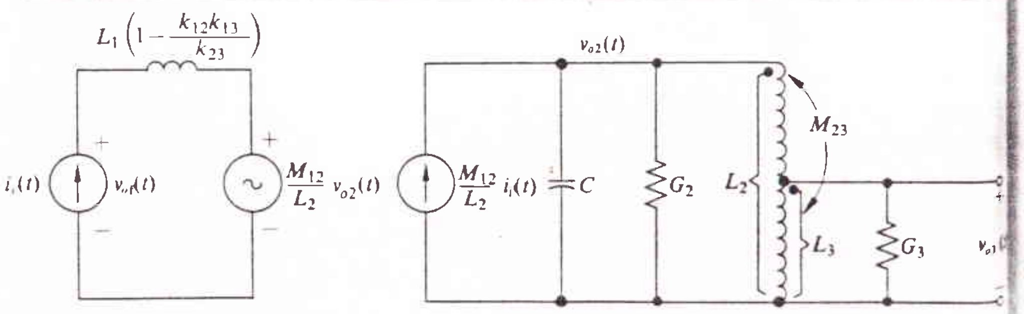

$L_1\left(1 - \frac{k_{12}k_{13}}{k_{23}}\right)$
$v_{o2}(t)$
$i_i(t)$
$v_{o1}(t)$
$\frac{M_{12}}{L_2}v_{o2}(t)$
$\frac{M_{12}}{L_2}i_i(t)$
C
G_2
L_2
M_{23}
L_3
G_3
v_{o1}

Fig. 2.6–3 Equivalent circuit for parallel resonant three-winding transformer.

---

† If Eq. (2.6–1) is not satisfied, Q_E is given by: $Q_E = \dfrac{1}{\omega_0 L_3(1 - k_{23}k_{13}/k_{13})G_3}$.

and

$$Q_T = \frac{\omega_0 C}{G_2 + G_3(M_{23}/L_2)^2},$$

then the circuit of Fig. 2.6–3 reduces to the final high-Q simplified form shown in Fig. 2.6–4. With the aid of this model we may immediately write

$$Z_{12}(p) = \frac{V_{o2}(p)}{I_i(p)} = \frac{M_{12}}{L_2}\frac{p\dfrac{1}{C}}{p^2 + \dfrac{pG_{eq}}{C} + \dfrac{1}{L_2C}}, \tag{2.6–4}$$

where $G_{eq} = G_2 + G_3(M_{23}/L_2)^2$. In addition, we obtain

$$Z_{13}(p) = \frac{V_{o3}(p)}{I_i(p)} = nZ_{12}(p) = \frac{M_{12}}{L_2}\frac{M_{23}}{L_2}\frac{p\dfrac{1}{C}}{p^2 + \dfrac{pG_{eq}}{C} + \dfrac{1}{L_2C}} \tag{2.6–5}$$

and

$$Z_{11}(p) = \frac{V_{o1}(p)}{I_i(p)} = pL_1\left(1 - \frac{k_{12}k_{13}}{k_{23}}\right) + \left(\frac{M_{12}}{L_2}\right)^2\frac{p\dfrac{1}{C}}{p^2 + \dfrac{pG_{eq}}{C} + \dfrac{1}{L_2C}}. \tag{2.6–6}$$

It is interesting to note that $Z_{12}(p)$ and $Z_{13}(p)$ are in the form of the impedance of a parallel RLC circuit multiplied by a scale factor, whereas $Z_{11}(p)$ contains an additional term which may be modeled as a series inductor. Consequently, if $i_i(t)$ were a periodic waveform (of period $2\pi/\omega_0$) containing several harmonics plus the fundamental, and Q_T were 10 or greater, $v_{o2}(t)$ and $v_{o3}(t)$ would be sinusoidal in form, whereas $v_{o1}(t)$ would be a periodic function with considerable harmonic content. This effect is due to the series inductor of $Z_{11}(p)$, which, unlike the parallel RLC circuit, does not appear as a small impedance at the harmonics of ω_0.

Example 2.6–1 For the circuit shown in Fig. 2.6–5 determine the value of M_{23} which maximizes $v_{o3}(t)$. Also determine (with the value of M_{23} found) an expression for $v_{o2}(t)$ and $v_{o3}(t)$ and values for Q_T and BW.

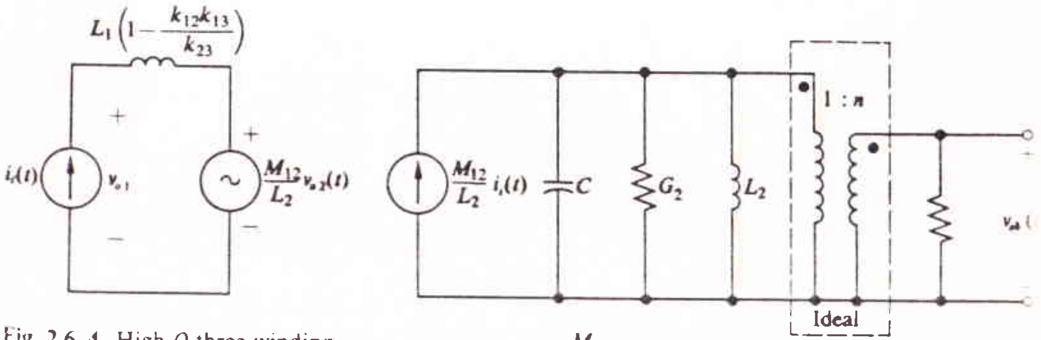

$L_1\left(1 - \frac{k_{12}k_{13}}{k_{23}}\right)$
$i_i(t)$
v_{o1}
$\frac{M_{12}}{L_2}v_{o2}(t)$
$\frac{M_{12}}{L_2}i_i(t)$
C
G_2
L_2
1 : n
Ideal
v_{o3}
$n = \frac{M_{23}}{L_2}$

Fig. 2.6–4 High-Q three-winding tuned transformer model.

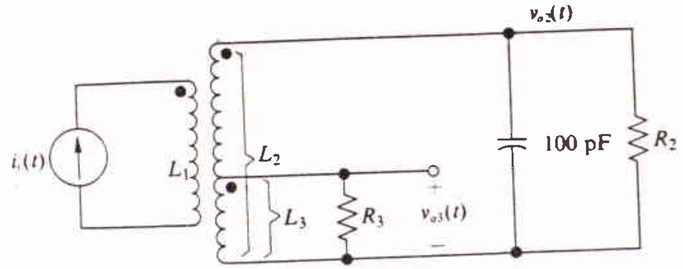


Figure 2.6–5

$i_s(t) = (1 \text{ mA}) \sin 10^7 t$

$R_2 = 400\ \text{k}\Omega$ $L_2 = 100\ \mu\text{H}$
$R_3 = 10\ \Omega$ $M_{12} = 10\ \mu\text{H}$
$L_1 = 2\ \mu\text{H}$ $k_{23} \approx 0.5$

Solution. If we assume $n'Q_{T'}Q_E > 100$, $Q_E > 10$, and $Q_T > 10$, we may employ the model of Fig. 2.6–4 omitting L_2 and C (which resonate in this case at $\omega_0 = 10^7$ rad/sec) as shown in Fig. 2.6–6. Now $v_{o3}(t)$ is maximized if n is chosen to match the 10 Ω resistor to the 400 kΩ resistor, i.e., if

$$n = \frac{M_{23}}{L_2} = \sqrt{\frac{10\ \Omega}{400\ \text{k}\Omega}} = 0.005.$$

Therefore, $M_{23} = 100\ \mu\text{H} \times 0.005 = \frac{1}{2}\ \mu\text{H}$, and, since $k_{23} = 0.5$, $L_3 = M_{23}^2/k_{23}^2 L_2 = 0.01\ \mu\text{H}$. Consequently,

$$Q_E = \frac{1}{\omega_0 L_3(1 - k_{23}^2)G_3} = \frac{400}{3}$$

and

$$Q_{T'} = \frac{\omega_0 C}{n^2}(10\ \Omega) \doteq 400,$$

and thus $n'Q_{T'}Q_E = k_{23}Q_{T'}Q_E = 80{,}000/3$.

To obtain an expression for v_{o3} we reflect the input current source and the 400-kΩ resistor across the 10 Ω output resistor. The reflected resistance is $(400\ \text{k}\Omega)n^2 = 10\ \Omega$ and the reflected current source is

$$\frac{M_{12}}{L_2}\frac{L_2}{M_{23}}(\sin 10^7 t)\ \text{mA} = 20\ \text{mA} \sin 10^7 t;$$

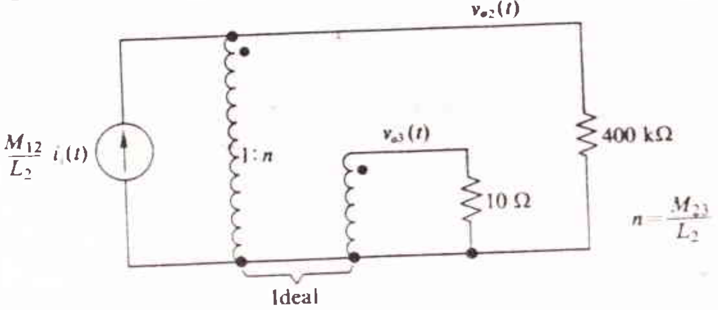


Figure 2.6–6

hence

$$v_{o3}(t) = (20\ \text{mA})(5\ \Omega) \sin 10^7 t = 0.1\ \text{V} \sin 10^7 t$$

and

$$v_{o2}(t) = \frac{v_{o3}(t)}{n} = 20\ \text{V} \sin 10^7 t.$$

Finally, by observing that R_3 reflects into the secondary as 400 kΩ, we obtain

$$Q_T = \frac{200\ \text{k}\Omega}{(10^7\ \text{rad/sec})(100\ \mu\text{H})} = 200$$

(a slightly higher value than one would normally expect from an inductor) and

$$\text{BW} = \frac{10^7\ \text{rad/sec}}{200} = 5 \times 10^4\ \text{rad/sec}.$$

Since $n'Q_{T'}Q_E \gg 100$, $Q_E > 10$, and $Q_T > 10$, use of the model is justified.

PROBLEMS

2.1 For the transformer with the loss terms shown in Figure 2.P–1, determine the value of L_2 required to match an 8 Ω load to a 3200 Ω source in the midband. What are the upper and lower −3 dB frequencies for this case? What is the open-circuit voltage transformation ratio of this transformer? What is the efficiency of the unit when working under the originally specified conditions?

2.2 Assume that the transformer (with the same values of k, L_1 and L_2) of Problem 2.1 is used between a 3 Ω load and a 1200 Ω source. Find the 3 dB frequencies and the efficiency.

2.3 Repeat Problem 2.2 for the case where the load is 20 Ω and the source is 8000 Ω.

2.4 Suppose the source output in Problem 2.1 is such that 10 mW of ac power reaches the 8 Ω load. Plot load power vs. load impedance for the case where the load varies between 2 Ω and 32 Ω while the other parameters of the system are maintained constant.

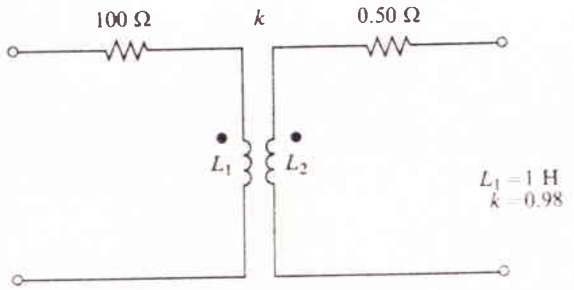


Figure 2.P–1

2.6

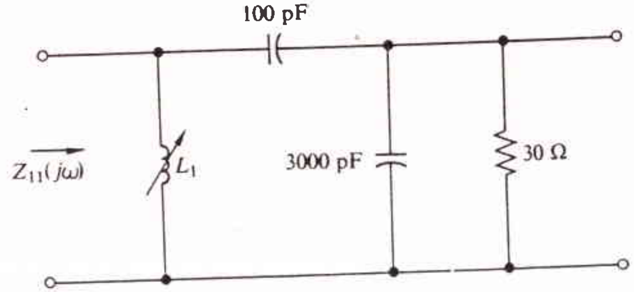


Figure 2.P–2

2.5 Consider the "matching network" of Figure 2.P–2, in which L_1 is tuned to make $Z_{11}(j\omega)$ resonate at $\omega_0 = 2 \times 10^6$ (L_1 has a Q of 100 at $\omega = 2 \times 10^6$ rps). Find $Z_{11}(j\omega_0)$ and the input Q_T. What is the approximate phase shift between $v_{o2}(t)$ and $v_{o1}(t)$ at $\omega = 2 \times 10^6$? If the peak sinusoid current flowing into $Z_{11}(j\omega)$ at resonance is 1 mA, what is the peak value of $v_{o2}(t)$?

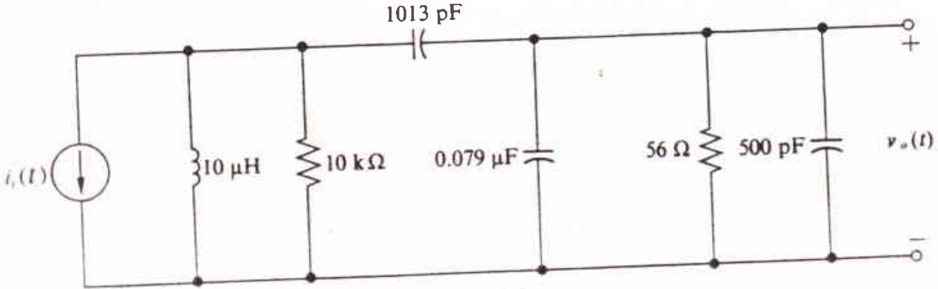
2.6 Repeat Problem 2.5 for the case where the 30 Ω load is increased to 1000 Ω.

2.7 A capacitor with a Q of 200 is combined with a coil with a Q of 80 and a 20 kΩ resistor to produce a parallel tuned circuit resonant at 60×10^6 rps. Letting $C = 25$ pF, find the bandwidth of the resultant circuit.

2.8 Repeat Problem 2.7 for the same coil and capacitor connected in series with a 10 Ω resistor to form a series resonant circuit at the same frequency. If the power input to this circuit at resonance is 10 W, how much power is dissipated in each series element?

2.9 For the circuit of Figure 2.P–3, determine $v_o(t)$ at resonance. If the capacitors are lossless and the coil loss is included in the 10 kΩ resistor, what is the circuit's Q_T? Assuming that the input current generator also produces a 0.5 mA peak sinusoidal component at both $2\omega_0$ and $3\omega_0$, estimate the relative distortion (each harmonic separately) at $v_o(t)$. (The pole-zero diagram of Fig. 2.4–2 provides an easy way to estimate distortion at harmonic frequencies.)

2.10 Assume that $i_1(t)$ in the circuit of Fig. 2.P–4 is a sinusoidal current with a peak value of 15 mA and a radian frequency of $\omega_0 = 4 \times 10^6$ rps and C_2 is chosen to resonate at ω_0.



$i_i(t) = (1 \text{ mA}) \sin \omega_0 t$

Figure 2.P–3

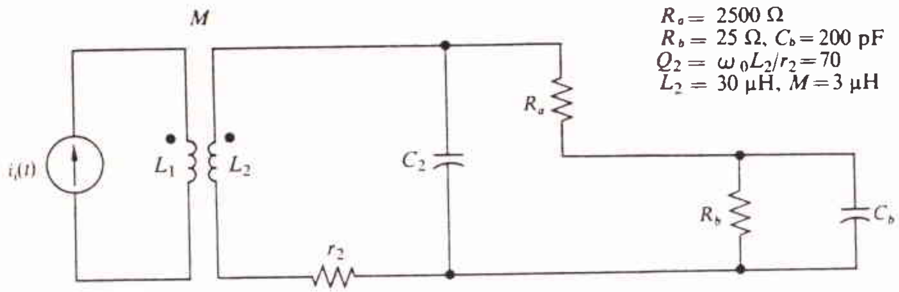

M
$R_a = 2500\ \Omega$
$R_b = 25\ \Omega,\ C_b = 200$ pF
$Q_2 = \omega_0 L_2/r_2 = 70$
$L_2 = 30\ \mu$H, $M = 3\ \mu$H
$i_i(t)$
L_1
L_2
C_2
R_a
R_b
C_b
r_2

Figure 2.P–4

Find the output voltage across R_b. Make reasonable assumptions. State these assumptions clearly. What is the Q of the circuit?

2.11 For each of the networks shown in Fig. 2.P–5, determine ω_0, Q_T, BW, $Z_{11}(j\omega_0)$, and $Z_{12}(j\omega_0)$.

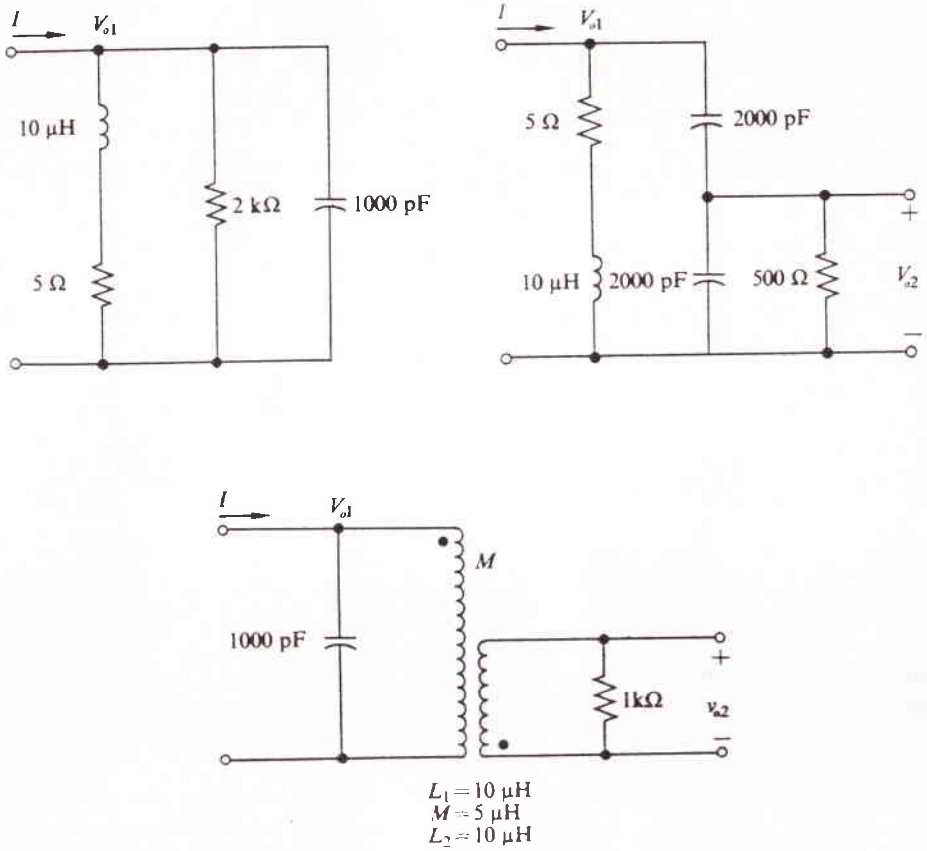


Figure 2.P–5

APPENDIX TO CHAPTER 2

TRANSFORMER EQUIVALENT CIRCUITS

The two-winding transformer shown in Fig. 2.A-1 may be described completely in terms of its terminal equations:

$$V_1(p) = pL_1I_1(p) + pMI_2(p),$$
$$V_2(p) = pMI_1(p) + pL_2I_2(p). \qquad (2.A\text{-}1)$$

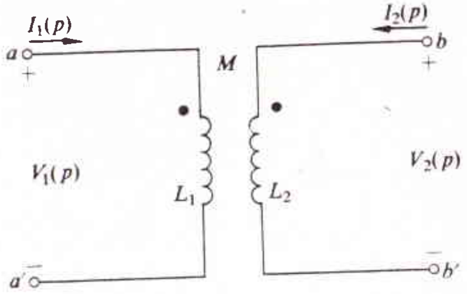


Fig. 2.A-1 Two-winding transformer.

Consequently, any network which has the same defining equations as the transformer may be substituted for the transformer in any circuit in which the transformer is placed, and this substitution will not affect the voltages or currents in the overall circuit. Several networks having the same defining equations as the transformer (and therefore referred to as terminal equivalents of the transformer) are shown in Fig. 2.A-2. By reflecting one or more of the various inductors through the ideal transformers in the circuits of Fig. 2.A-2, one can obtain a number of other terminal equivalent networks for the transformer.

There are several procedures for verifying that the networks of Fig. 2.A-2 are indeed terminal equivalents for the transformer or for evaluating the equivalent inductances of the new networks. One approach is to observe that Eq. (2.A-1) is equivalent to a statement that the impedance at *a–a'* with *b–b'* open is pL_1, that the impedance at *b–b'* with *a–a'* open is pL_2, and that the ratio of the voltage at *b–b'* to a current applied at *a–a'* with *b–b'* open is pM. We then choose the values of the proposed equivalent circuit so that these three conditions are met. The values shown in Fig. 2.A-2 do meet these conditions.

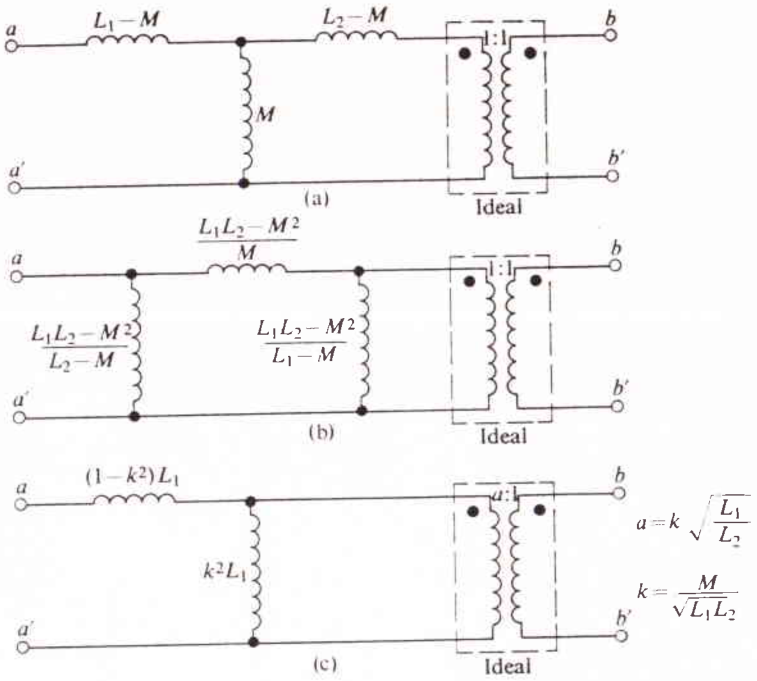
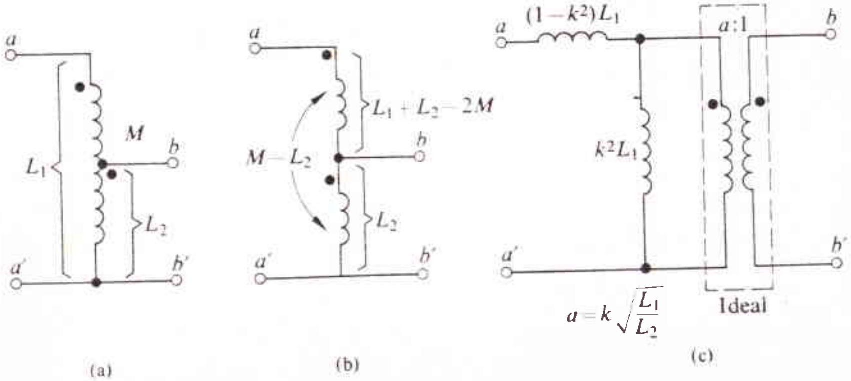

a
$L_1 - M$
$L_2 - M$
1:1
b
M
a'
b'
(a)
Ideal
a
$\frac{L_1L_2 - M^2}{M}$
1:1
b
$\frac{L_1L_2 - M^2}{L_2 - M}$
$\frac{L_1L_2 - M^2}{L_1 - M}$
a'
b'
(b)
Ideal
a
$(1-k^2)L_1$
a:1
b
$a = k\sqrt{\frac{L_1}{L_2}}$
k^2L_1
$k = \frac{M}{\sqrt{L_1L_2}}$
a'
b'
(c)
Ideal

Fig. 2.A-2 Terminal equivalent networks for transformers.



(a) (b) (c)

Fig. 2.A-3 Terminal equivalent transformers.

A transformer may have a terminal equivalent network which is also a transformer. To illustrate this, Fig. 2.A-3 shows two of the many terminal equivalent networks for an auto transformer (M is the mutual inductance between L_1 and L_2).

Terminal equivalent networks for the transformer of Fig. 2.A-1 may also take the form of networks containing controlled sources. Three such networks are shown in Fig. 2.A-4.

The dots at the ends of the transformer windings in Figs. 2.A–1, 2.A–2, and 2.A–3 show the relative directions of induced voltages. A current flowing into a dotted winding causes a plus-to-minus drop across this winding and induces a voltage in all other windings such that the dotted end is positive. Moving the dot to the opposite end of the transformer winding on the secondary of Fig. 2.A–1 would reverse the signs of both of the pM-terms in Eq. (2.A–1). It would also lead to a reversal of the secondary dot on the transformers of Fig. 2.A–2 and to the reversal of the direction of both of the generators in Figs. 2.A–4(a), (b), and (c). For the auto transformer, placing the dot at the bottom of L_2 would reverse the sign of the mutual term in Fig. 2.A–3(b) and reverse the dot on the secondary of the transformer of Fig. 2.A–3(c).

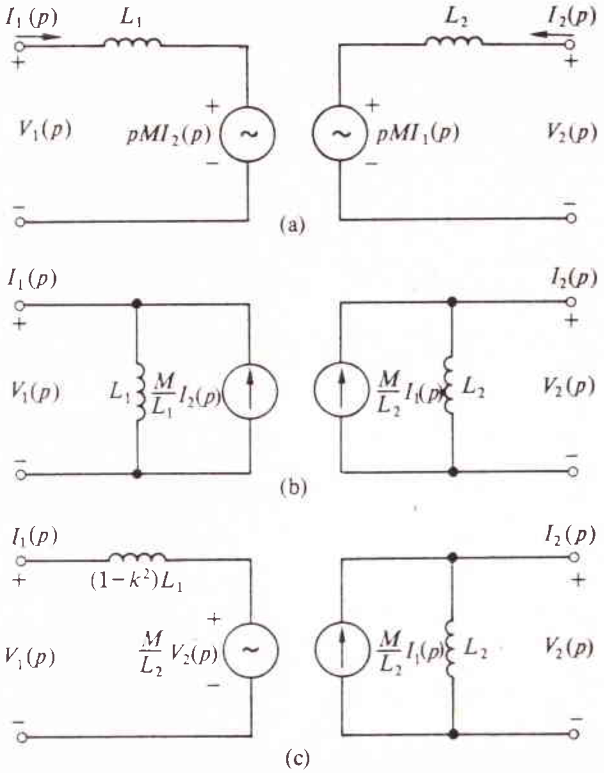


Fig. 2.A–4 Terminal equivalent networks containing controlled sources.

CHAPTER 3

TRANSMISSION OF SIGNALS THROUGH NARROWBAND FILTERS

The purpose of this chapter is to examine a simplified method of determining the output response of a class of "narrowband" high-frequency networks when these networks are driven by any one of a variety of useful signals. Among these signals we include steps, impulses, and various amplitude-modulated signals centered within the passband of the filter.

First we define the class of networks to be considered, then we define the low-pass equivalent circuit for such networks, and finally we show that the response of the original network to a given signal is a function of the response of the low-pass equivalent circuit either to the signal or, in the case of amplitude-modulated signals, to the envelope of the modulated signal. Thus, where the method is applicable, the original complicated network response problem is replaced by a simplified and reasonably accurate approximate solution.

It has been suggested to us that such material is covered elsewhere and need not be repeated here. Our experience has been that at least a brief review of these concepts, perhaps from a viewpoint that the reader has not encountered before, is useful in understanding their use in later chapters.

3.1 LOW-PASS EQUIVALENT NETWORKS FOR SYMMETRICAL BANDPASS NETWORKS

For a general narrowband network whose transfer function is $H(p)$ and whose passband is centered about ω_0, the magnitude and phase of $H(j\omega)$ vs. ω take the form shown in Fig. 3.1–1. If $H(j\omega)$ is indeed narrowband, then the amplitude response, $|H(j\omega)|$, falls essentially to zero a short distance on either side of ω_0, as shown. In

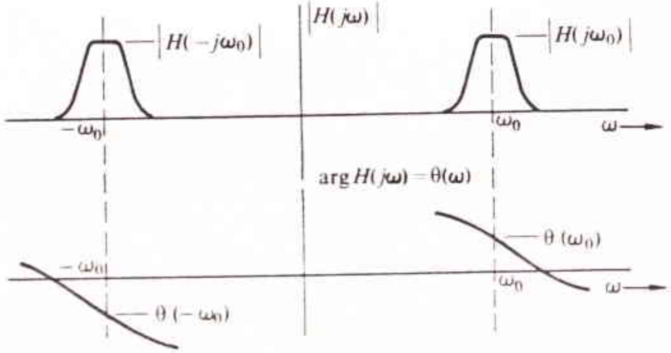


Fig. 3.1–1 Magnitude and phase plot for a narrowband network.

ddition, if $H(j\omega)$ represents the transfer function of a physical network with a real not imaginary or complex) impulse response $h(t)$, then $H(-j\omega) = H^*(j\omega)$,† which is quivalent to

$$|H(-j\omega)| = |H(j\omega)| \tag{3.1–1}$$

nd

$$\theta(-\omega) = -\theta(\omega),$$

vhere $\theta(\omega) \equiv \arg H(j\omega)$. Equation (3.1–1) is, of course, simply the statement that the nagnitude of a physical network must be an even function of ω and the phase must e an odd function of ω.

We define the low-pass equivalent transfer function $H_L(j\omega)$ for $H(j\omega)$ as that unction whose magnitude has the same dependence on ω in the vicinity of $\omega = 0$ as $H(j\omega)|$ has in the vicinity of $\omega = \omega_0$, and whose phase angle has the same dependence n ω in the vicinity of $\omega = 0$ as $\theta(\omega) - \theta(\omega_0)$ has in the vicinity of ω_0. This definition nay also be expressed in the form

$$H_L(j\omega) \equiv H(j\omega + j\omega_0)e^{-j\theta(\omega_0)}u(\omega + \omega_0), \tag{3.1–2}$$

vhere $u(\omega + \omega_0)$ is the unit step function which, in Eq. (3.1–2), has the effect of removing the lower portion ($\omega < 0$) of $H(j\omega)$. The $e^{-j\theta(\omega_0)}$ term removes a constant phase angle $\theta(\omega_0)$ from the phase of $H_L(j\omega)$. A plot of the magnitude and phase of $H_L(j\omega)$, which correspond to the magnitude and phase of $H(j\omega)$ shown in Fig. 3.1–1, appears in Fig. 3.1–2. Note that the phase of $H_L(j\omega)$ is zero for $\omega = 0$.

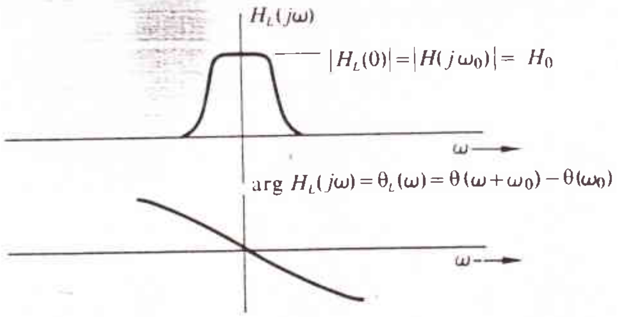


Fig. 3.1–2 Magnitude and phase plot for $H_L(j\omega)$.

The transfer function $H(j\omega)$ is defined as being symmetric about ω_0 if $H_L(-j\omega) = H_L^*(j\omega)$, or equivalently

$$|H_L(-j\omega)| = |H_L(j\omega)| \tag{3.1–3}$$

and

$$\theta_L(\omega) = -\theta_L(-\omega).$$

In this chapter we shall restrict our attention to symmetric narrowband networks for two reasons. First and most important, most physical narrowband networks may be closely approximated by symmetric transfer functions. Second, the symmetry condition, expressed by Eq. (3.1–3), ensures that the impulse response $h_L(t)$ of the low-pass equivalent network is real, which is essential if $H_L(j\omega)$ is to be physically realizable. If $H(j\omega)$ is not symmetric about ω_0, the low-pass equivalent transfer function is still defined by Eq. (3.1–2); however, it may not be associated with a physical network and, in addition, calculations employing it may become somewhat more involved.

---

† By definition, $H(j\omega) = \int_{-\infty}^{\infty} h(t)e^{-j\omega t}\,dt$. If the impulse response $h(t)$ is real, then

$$H(-j\omega) = \int_{-\infty}^{\infty} h(t)e^{-(-j\omega)t}\,dt = H^*(j\omega),$$

where $H^*(j\omega)$ is the complex conjugate of $H(j\omega)$.

In addition to relating $H_L(j\omega)$ to $H(j\omega)$, for the symmetric narrowband filter we may obtain the inverse relationship. Specifically, with the aid of Eq. (3.1–1) we note that

$$H(j\omega) = H_L(j\omega - j\omega_0)e^{j\theta(\omega_0)} + H_L(j\omega_0 + j\omega)e^{-j\theta(\omega_0)}; \tag{3.1–4}$$

this equation, of course, represents $H_L(j\omega)$, with the appropriate phase angle added, shifted up and down in frequency by an amount ω_0.

To solidify our ideas on the relationship between $H_L(j\omega)$ and $H(j\omega)$ we shall develop low-pass equivalent networks for several physical narrowband networks. We consider first the parallel RLC circuit shown in Fig. 3.1–3. From Section 2.1 we recall that at resonance $Z_{11}(j\omega_0)$ is purely resistive; hence $\theta(\omega_0) = 0$. In addition, we recall that with $Q_T = \omega_0 RC > 10$, which is a necessary condition for a narrow bandwidth, $Z_{11}(j\omega)$ is closely approximated by (cf. Eq. 2.2–9)

$$Z_{11}(j\omega) = \frac{R}{1 + j\dfrac{\omega - \omega_0}{\alpha}}, \qquad \omega > 0,$$

where $\omega_0 = 1/\sqrt{LC}$ and $\alpha = 1/2RC$. Consequently, with $\theta(\omega_0) = 0$, Eq. (3.1–2) yields

$$Z_{11_L}(j\omega) = \frac{R}{1 + j\omega/\alpha}. \tag{3.1–5}$$

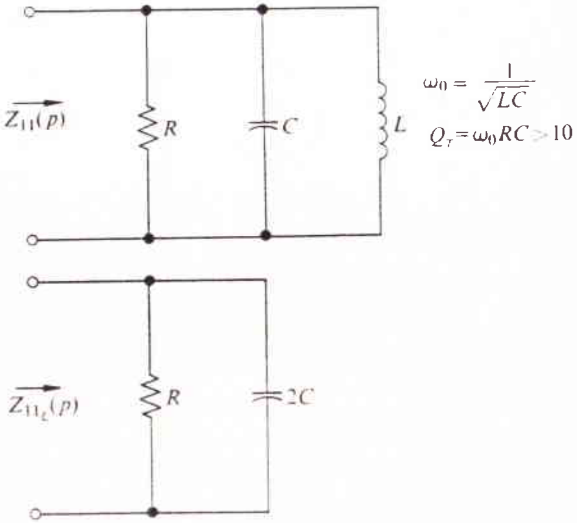

$\overrightarrow{Z_{11}(p)}$
R
C
L
$\omega_0 = \dfrac{1}{\sqrt{LC}}$
$Q_T = \omega_0 RC > 10$
$\overrightarrow{Z_{11_L}(p)}$
R
$2C$

Fig. 3.1–3 Parallel RLC circuit and its low-pass equivalent.

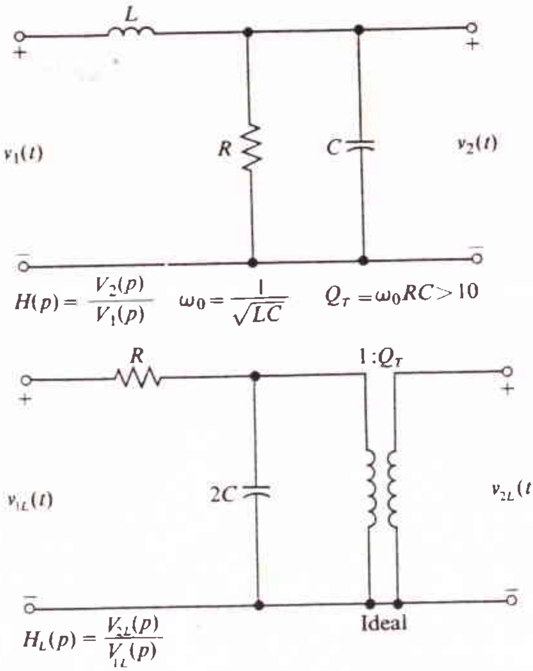

L
+
$v_1(t)$
R
C
$v_2(t)$
$H(p)=\frac{V_2(p)}{V_1(p)}$
$\omega_0=\frac{1}{\sqrt{LC}}$
$Q_T=\omega_0 RC>10$
R
$1:Q_T$
$v_{1L}(t)$
2C
$v_{2L}(t)$
Ideal
$H_L(p)=\frac{V_{2L}(p)}{V_{1L}(p)}$

Fig. 3.1-4 Narrowband circuit and its low-pass equivalent.

which we recognize as the input impedance of the low-pass equivalent network also shown in Fig. 3.1-3.

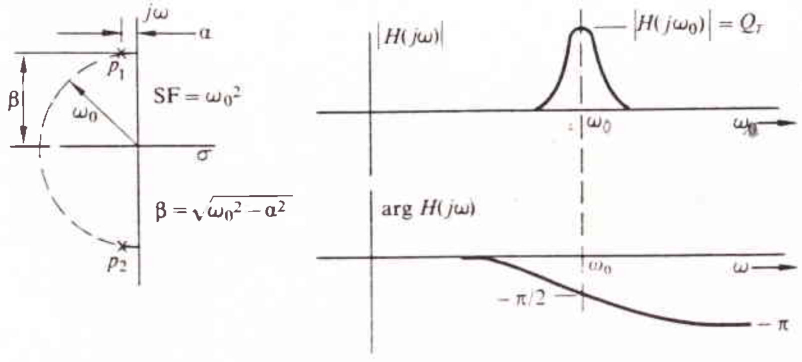


Fig. 3.1-5 Pole zero diagram of $H(p)$ and plot of $|H(j\omega)|$ and arg $(Hj\omega)$ vs. ω.

As a second example we consider the circuit shown in Fig. 3.1-4. For this circuit the voltage transfer function $H(p)$ is given by

$$H(p) = \frac{V_2(p)}{V_1(p)} = \frac{\omega_0^2}{p^2 + 2\alpha p + \omega_0^2}, \tag{3.1-6}$$

where $\omega_0 = 1/\sqrt{LC}$ and $\alpha = 1/2RC$. A pole-zero diagram of $H(p)$ and a sketch of

$|H(j\omega)|$ and arg $H(j\omega)$, for the case where $\omega_0/2\alpha = \omega_0 RC = Q_T > 10$, are shown in Fig. 3.1-5. When the complex poles are close to the imaginary axis (a necessary condition for the network to have a narrow bandwidth), then $H(j\omega)$ is significantly different from zero only in the vicinity of $\pm\omega_0$. Therefore, when graphically evaluating $H(j\omega)$ for $\omega > 0$ from the pole-zero diagram, we can closely approximate the phasor ρ_{p_2} drawn from p_2 to the imaginary axis by

$$\rho_{p_2} \approx 2\omega_0 e^{j(\pi/2)} \tag{3.1-7}$$

over the significant frequency range in the vicinity of ω_0. In addition, the phasor from p_1 to a point ω on the imaginary axis may be written as

$$\rho_{p_1} = \alpha + j(\omega - \beta) \approx \alpha + j(\omega - \omega_0). \tag{3.1-8}$$

Consequently, for $\omega > 0$ and $Q_T > 10$, $H(j\omega)$ may be closely approximated by

$$H(j\omega) = \frac{\text{SF}}{\rho_{p_1}\rho_{p_2}} = \frac{\omega_0^2 e^{-j(\pi/2)}}{2\omega_0\alpha\left(1 + j\dfrac{\omega - \omega_0}{\alpha}\right)} = \frac{Q_T e^{-j(\pi/2)}}{1 + j\dfrac{\omega - \omega_0}{\alpha}}. \tag{3.1-9}$$

If we now note that $\theta(\omega_0) = -\pi/2$ and employ Eq. (3.1-2), we obtain

$$H_L(j\omega) = \frac{Q_T}{1 + j(\omega/\alpha)}. \tag{3.1-10}$$

A non-unique low-pass equivalent network having this transfer function is shown also in Fig. 3.1-4. The ideal transformer is required to provide the voltage amplification of Q_T at $\omega = 0$, which is the amplification of the bandpass network at $\omega = \omega_0$. The bandwidth BW of $H(j\omega)$ is 2α, which is exactly twice the bandwidth of $H_L(j\omega)$.

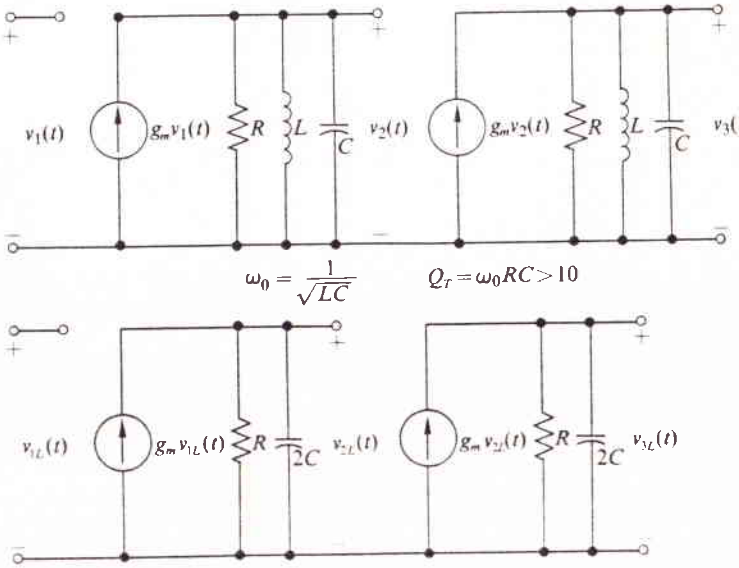


Fig. 3.1-6 Model for two-stage IF strip and its low-pass equivalent.

If two or more symmetric noninteracting narrowband filters, each of which has a center frequency of ω_0, are connected in cascade, then the low-pass equivalent of the composite filter is obtained by cascading the low-pass equivalents of the individual filters. Figure 3.1–6 illustrates the low-pass equivalent network comprising two cascaded, noninteracting parallel *RLC* circuits.

3.2 IMPULSE AND STEP RESPONSE

With the aid of the results of Section 3.1 we can represent the impulse and step response for a symmetric narrowband filter as a function of the impulse response $h_L(t)$ of the equivalent low-pass filter. Such a representation permits one to analyze the much simpler low-pass equivalent filter to obtain corresponding results for the bandpass filter.

If we designate $h(t)$ as the impulse response of the narrowband filter whose transfer function is given by $H(j\omega)$, then $h(t)$ is the inverse Fourier transform of $H(j\omega)$, that is,

$$h(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(j\omega)e^{j\omega t}\,d\omega. \tag{3.2–1}$$

With the aid of Eq. (3.1–4) we can express $H(j\omega)$ in terms of its low-pass equivalent $H_L(j\omega)$ to obtain

$$\begin{aligned} h(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H_L(j\omega - j\omega_0)e^{j\theta(\omega_0)}e^{j\omega t}\,d\omega \\ &\quad + \frac{1}{2\pi}\int_{-\infty}^{\infty} H_L(j\omega_0 + j\omega)e^{-j\theta(\omega_0)}e^{j\omega t}\,d\omega. \end{aligned} \tag{3.2–2}$$

If, in addition, we substitute $\omega' = \omega - \omega_0$ in the first integral and $\omega' = \omega_0 + \omega$ in the second integral of Eq. (3.2–2), the expression for $h(t)$ simplifies to the desired form

$$\begin{aligned} h(t) &= \left\{e^{j[\omega_0 t + \theta(\omega_0)]} + e^{-j[\omega_0 t + \theta(\omega_0)]}\right\}\frac{1}{2\pi}\int_{-\infty}^{\infty} H_L(j\omega')e^{j\omega' t}\,d\omega' \\ &= 2h_L(t)\cos[\omega_0 t + \theta(\omega_0)], \end{aligned} \tag{3.2–3}$$

respuesta a pulso

which directly relates $h(t)$ to $h_L(t)$, the impulse response of the low-pass equivalent circuit.

As an application of this result, let us evaluate the impulse response $z_{11}(t)$ of the high-Q parallel *RLC* circuit shown in Fig. 3.1–3. For this circuit $\theta(\omega_0) = 0$. The impulse response of the low-pass equivalent circuit $z_{11_L}(t)$ (also shown in Fig. 3.1–3) is known to be

$$z_{11_L}(t) = \frac{1}{2C}e^{-\alpha t}u(t), \tag{3.2–4}$$

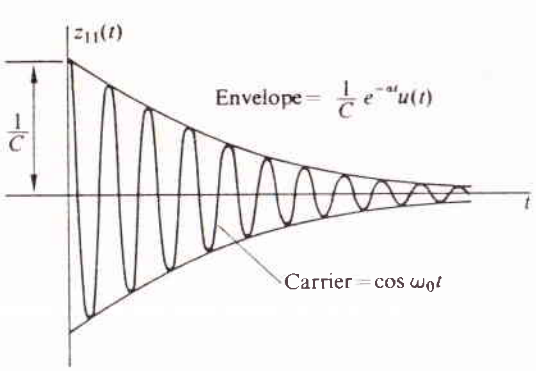


Fig. 3.2–1 Sketch of $z_{11}(t)$ for a parallel *RLC* circuit.

where $\alpha = 1/2RC$. Consequently, by employing Eq. (3.2–3) we obtain

$$z_{11}(t) = \frac{1}{C}e^{-\alpha t}(\cos\omega_0 t)u(t), \tag{3.2–5}$$

which is exactly the expression obtained for $z_{11}(t)$ in Eq. (2.2–12) with more conventional techniques. A sketch of $z_{11}(t)$ is shown in Fig. 3.2–1.

To obtain the step response of $a(t)$ for a narrowband filter in terms of its low-pass equivalent circuit, we first evaluate the step response in terms of $H(j\omega)$, then express $H(j\omega)$ in terms of $H_L(j\omega)$, and finally find $a(t)$ in terms of $h_L(t)$.

The step response $a(t)$ for the narrowband filter can be written in the form

$$a(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} H(j\omega)\left[\frac{1}{j\omega} + \pi\delta(\omega)\right]e^{j\omega t}\,d\omega, \tag{3.2–6}$$

where $(1/j\omega) + \pi\delta(\omega)$ is the Fourier transform of the unit step. Since for the networks under consideration the response at dc is always assumed to be zero, it follows that $H(0) = 0$ and therefore the $\delta(\omega)$-term may be omitted from Eq. (3.2–6).

We now express $H(j\omega)$ in terms of $H_L(j\omega)$ and restrict our attention to regions close enough to the complex poles so that near the upper complex pole $j\omega$ may be replaced by $j\omega_0$ and near the lower complex pole $j\omega$ may be replaced by $-j\omega_0$. After some rearrangement, one obtains the step response in terms of the product of the impulse response of the low-pass equivalent circuit and a sine wave at the center frequency of the narrowband circuit; that is,

respuesta al escalon

$$a(t) = +\frac{2h_L(t)}{\omega_0}\sin[\omega_0 t + \theta(\omega_0)]. \tag{3.2–7}$$

Evaluating Eq. (3.2–7) for a high-Q parallel *RLC* circuit, we obtain

$$a(t) = \frac{1}{\omega_0 C}e^{-\alpha t}(\sin\omega_0 t)u(t).$$

where $\omega_0 = 1/\sqrt{LC}$ and $\alpha = 1/2RC$.

Physically what this means is that driving the tuned circuit with a step function causes it to "ring" at its resonant frequency. Since there is no continuing supply of energy at this resonant frequency, the "ringing" decays with time. The higher the Q of the circuit, the more cycles it takes for the decay to fall to any given percentage of the original level.

3.3 NARROWBAND NETWORKS WITH MODULATED INPUTS

In this section we shall apply an amplitude-modulated signal of the form $s_i(t) = g(t) \cos \omega_0 t$, shown in Fig. 3.3–1, to the symmetric narrowband filter and demonstrate that the filter output is of the form

$s_i(t) = g(t) \cos w_0 t$

$$s_0(t) = [g(t) * h_L(t)] \cos [\omega_0 t + \theta(\omega_0)],$$

where $*$ denotes convolution. This result is quite significant, since it simplifies the problem of calculating the response of a bandpass filter excited by an AM wave to the problem of calculating the response of the equivalent low-pass filter excited by the envelope waveform $g(t)$. The resultant expression provides the modulation for the output carrier $\cos [\omega_0 t + \theta(\omega_0)]$.

To begin our development we define $G(\omega)$ as the Fourier transform of $g(t)$ and assume that $|G(\omega)|$ has some form similar to that shown in Fig. 3.3–2. With this definition of $G(\omega)$ and with the aid of the "shifting theorem," the Fourier transform of

$$s_i(t) = \frac{g(t)}{2}(e^{j\omega_0 t} + e^{-j\omega_0 t})$$

takes the form

$$S_i(\omega) = \frac{G(\omega + \omega_0)}{2} + \frac{G(\omega - \omega_0)}{2},$$

which is also shown in Fig. 3.3–2. We assume throughout this analysis that the two terms in the expression for $S_i(\omega)$ do not overlap. This assumption requires that the highest-frequency component ω_m of $g(t)$ be lower than the carrier frequency ω_0.

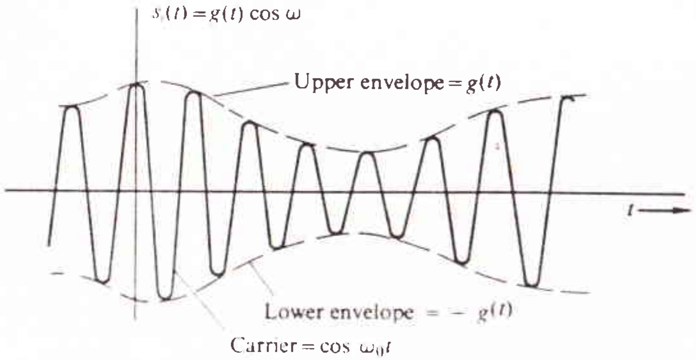


Fig. 3.3–1 Typical plot of amplitude-modulated waveform.

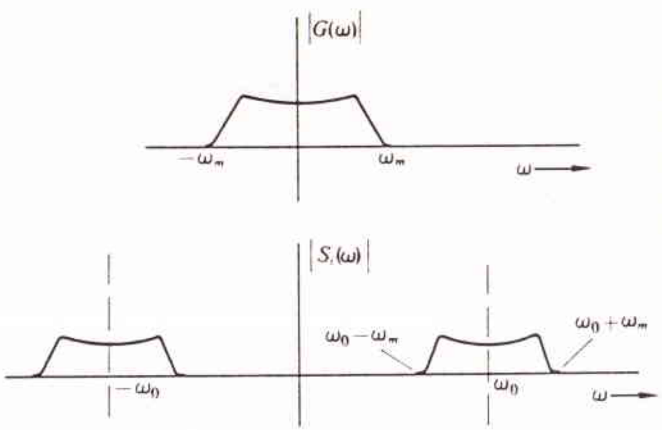


Fig. 3.3–2 Plot of $|G(\omega)|$ and $|S_i(\omega)|$ vs. ω.

This condition is almost always satisfied in practice, since for most AM systems ω_m is required to be several orders of magnitude smaller than ω_0 because of the physical limitations of the AM modulator and demodulator.

For the condition $\omega_0 \gg \omega_m$, we see that the spectrum of $s_i(t)$ occupies a narrow band of frequencies centered on the carrier frequency ω_0. This property permits frequency division multiplexing (FDM), i.e., the independent combination of many AM signals within a single channel. This is accomplished by choosing a distinct carrier frequency for each signal in such a fashion that none of the signal spectra overlap in the frequency domain. To extract a desired signal at the receiving end of the channel, one need only pass the composite signal through a narrowband filter: the filter passes the signal of interest and attenuates all others.

The filter capable of passing the desired AM signal centered at ω_0 would, of course, have the form shown in Fig. 3.1–1. With $s_i(t)$ applied to this filter, whose transfer function is $H(j\omega)$, the output signal $s_o(t)$ may be written in the form

$$s_o(t) = \mathscr{F}^{-1}[H(j\omega)S_i(\omega)], \qquad (3.3\text{–}1)$$

where $\mathscr{F}^{-1}$ indicates the inverse Fourier transform operation. With the aid of Eq. (3.1–4), which relates $H(j\omega)$ to its low-pass equivalent, the expression for $s_o(t)$ may be rewritten as

$$s_o(t) = \mathscr{F}^{-1}\left\{[H_L(j\omega - j\omega_0)e^{j\theta(\omega_0)} + H_L(j\omega_0 + j\omega)e^{-j\theta(\omega_0)}]\left[\frac{G(\omega + \omega_0)}{2} + \frac{G(\omega - \omega_0)}{2}\right]\right\}$$

$$= \frac{e^{j\theta(\omega_0)}}{2}\mathscr{F}^{-1}[H_L(j\omega - j\omega_0)G(\omega - \omega_0)] + \frac{e^{-j\theta(\omega_0)}}{2}\mathscr{F}^{-1}[H_L(j\omega + j\omega_0)G(\omega + \omega_0)]. \qquad (3.3\text{–}2)$$

Equation (3.3–2) makes use of the fact that the upper portion of the spectrum of $S_i(\omega)$ does not overlap the lower portion of the spectrum of $H(j\omega)$ and vice versa.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA

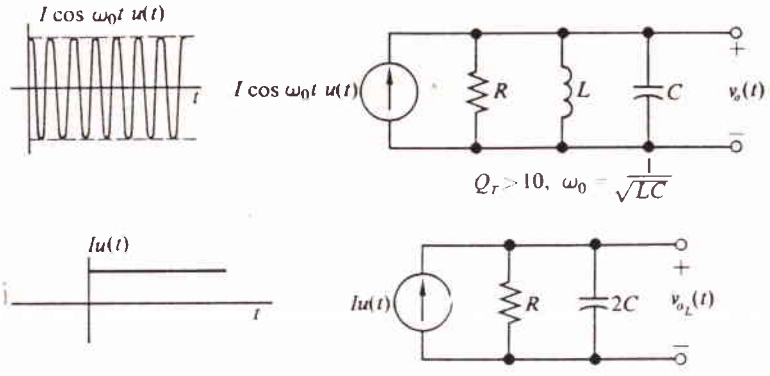


Fig. 3.3–3 Parallel RLC circuit with AM input.

If the inverse shifting theorem is now employed, Eq. (3.3–2) takes the desired form

$$s_o(t) = \frac{e^{j\omega_0 t}e^{j\theta(\omega_0)} + e^{-j\omega_0 t}e^{-j\theta(\omega_0)}}{2}\mathscr{F}^{-1}[H_L(j\omega)G(\omega)]$$

$$= [g(t) * h_L(t)]\cos[\omega_0 t + \theta(\omega_0)]. \tag{3.3–3}$$

Note carefully that $\mathscr{F}^{-1}[H_L(j\omega)G(\omega)] = g(t) * h_L(t)$ is simply a shorthand mathematical expression, which we refer to as the "convolution of $h_L(t)$ with $g(t)$." Although this convolution can be determined with the aid of the convolution integral, it is in general obtained in a more straightforward fashion as the output of the network whose impulse response is $h_L(t)$ and whose input signal is $g(t)$, that is, as the output of the low-pass equivalent network excited by $g(t)$. A few examples should clarify this point.

As a first example, let us evaluate the voltage across a high-Q_T parallel RLC circuit in which a current of the form $I\cos\omega_0 t$ is applied at $t = 0$. The narrowband network shown in Fig. 3.3–3 is being excited by an AM wave with an envelope $g(t) = Iu(t)$; therefore, Eq. (3.3–3) may be employed to determine $v_o(t)$. As was pointed out above, $g(t) * h_L(t) = v_{o_L}(t)$ is the output of the low-pass equivalent of the parallel RLC circuit with a current $Iu(t)$ applied at the input. The step response of the low-pass equivalent circuit is readily found to be

$$v_{o_L}(t) = IR(1 - e^{-\alpha t})u(t);$$

and since $\theta(\omega_0) = 0$, Eq. (3.3–3) (where $\alpha = 1/2RC$ and $\omega_0 = 1/\sqrt{LC}$) yields

$$v_o(t) = IR(1 - e^{-\alpha t})\cos\omega_o t u(t). \tag{3.3–4}$$

A sketch of $v_{o_L}(t)$ and $v_o(t)$ is given in Fig. 3.3–4. We observe that the steady-state value of $v_o(t)$ is the product of the input current $I\cos\omega_0 t$ and the resistance R of the parallel RLC circuit. This is an expected result because at the resonant frequency $Z_{11}(j\omega_0) = R$. In addition, we observe that $v_o(t)$ rises toward its steady-state value with an envelope governed by a single time constant $\tau = 1/\alpha$; thus for $t \geq 4\tau$, $v_o(t)$ has attained an amplitude within 2% of its steady-state value.

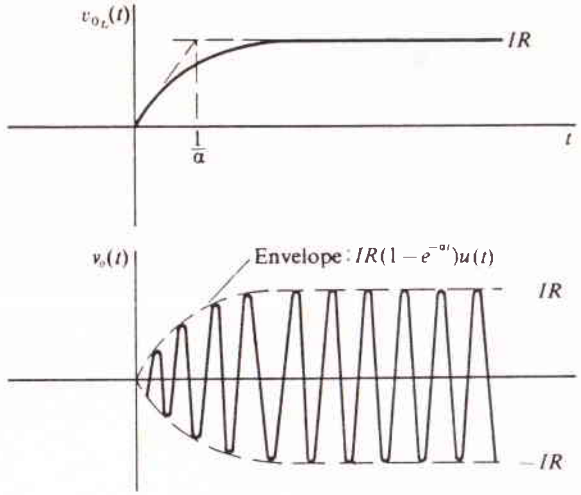


Fig. 3.3–4 Output waveforms for the circuits of Fig. 3.3–3.

It is apparent that the closer the complex poles of the parallel RLC circuit lie to the imaginary axis in the complex p-plane, the longer it takes the output waveform to reach steady state. Consequently, networks with very narrow bandwidths ($\alpha \ll \omega_0$) are capable of transmitting without distortion only input AM waves with slowly varying envelopes, or equivalently, AM waves whose spectra are contained within the passband of the narrowband filter. An alternative interpretation is that an AM wave is transmitted without distortion if the spectrum of $g(t)$ lies within the passband of $H_L(j\omega)$.

As a second example, let us apply a periodically gated carrier of frequency ω_0 to a high-Q parallel RLC circuit as shown in Fig. 3.3–5. The input current may be

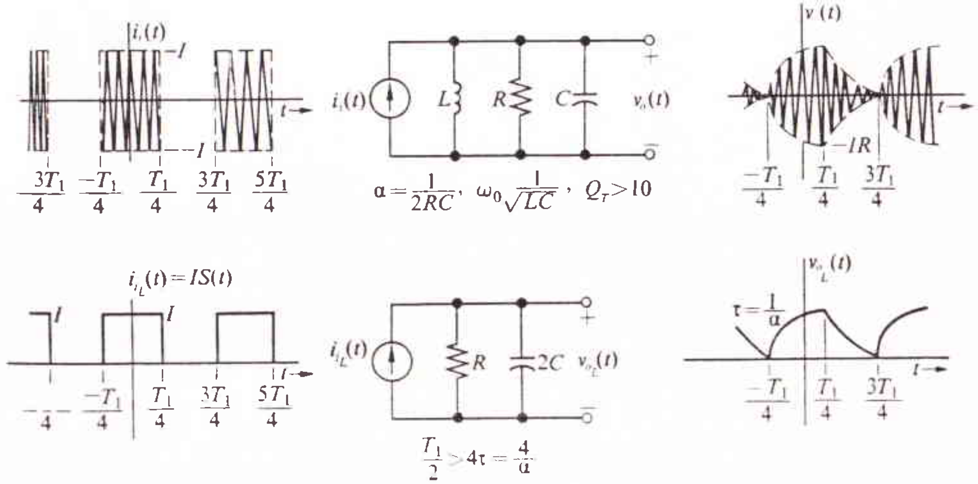


Fig. 3.3–5 Parallel RLC circuit driven by gated carrier.

resented as

$$i_i(t) = IS(t)\cos\omega_0 t,$$

ere $S(t)$ is a periodic switching function of period T_1 that has a value of either r 0. With this representation it is apparent that $i_i(t)$ is an AM wave and therefore t

$$v_o(t) = v_{o_L}(t)\cos\left[\omega_0 t + \theta(\omega_0)\right],$$

ere $v_{o_L}(t) = IS(t) * h_L(t)$ is the output of the low-pass equivalent of the parallel .C circuit driven by $i_{i_L}(t) = IS(t)$. Figure 3.3–5 also illustrates the low-pass uivalent circuit for determining $v_{o_L}(t)$.

If we assume that $T_1/2 > 4\tau = 4/\alpha$ (i.e., that the circuit reaches steady state in ch interval of duration $T_1/2$), then we can write $v_{o_L}(t)$ in the form

$$\begin{aligned} v_{o_L}(t) &= IR[1 - e^{-\alpha(t+T_1/4)}], && -T_1/4 \le t < T_1/4,\\ &= IRe^{-\alpha(t-T_1/4)}, && T_1/4 \le t < 3T_1/4,\\ &= IR[1 - e^{-\alpha(t-3T_1/4)}], && 3T_i/4 \le t < 5T_1/4, \qquad (3.3\text{–}5)\\ &\ \ \vdots && \quad\vdots \end{aligned}$$

nd in turn $v_o(t) = v_{o_L}(t)\cos\omega_0 t$. A sketch of $v_{o_L}(t)$ and $v_o(t)$ appears in Fig. 3.3–5 to e right of the corresponding circuits. Note that an output exists for the parallel RLC rcuit during the intervals of time during which there is no input. This phenomenon is early the result of the oscillatory decay of a high-Q circuit which has acquired nergy from an input signal during some previous interval of time.

As a third example, let us apply a sinusoidally modulated AM signal of the form

$$s_i(t) = A[1 + m\cos\omega_m t]\cos\omega_0 t$$

a general symmetric narrowband network whose passband is centered about ω_0. igure 3.3–6 illustrates the waveform of $s_i(t)$. The parameter m for this form of AM

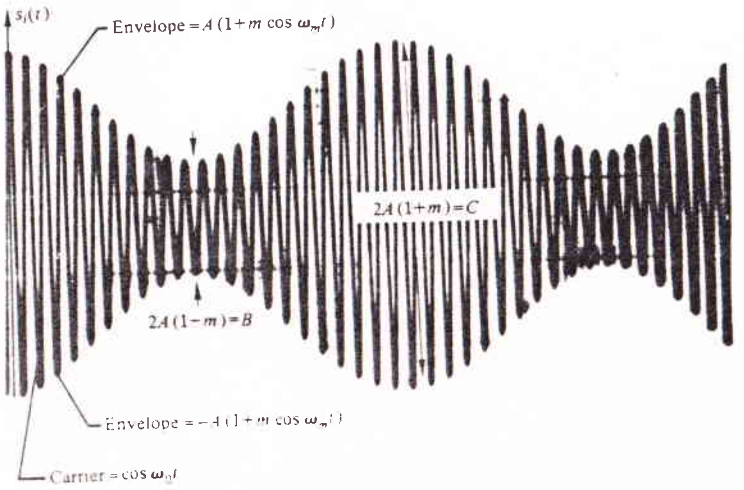


Fig. 3.3–6 Plot of sinusoidally modulated AM wave.

wave is defined as the modulation index, and A is defined as the unmodulated carrier amplitude. The modulation index m may be related to the waveform $s_i(t)$ by the relationship

$$m = \frac{C - B}{C + B}, \qquad (3.3\text{–}6)$$

where C is the maximum peak-to-peak value of $s_i(t)$ and B is the minimum peak-to-peak value of $s_i(t)$. If $B = 0$, then $m = 1$ and $s_i(t)$ is said to be 100% modulated. In general the percent modulation is given by

$$\%\ \text{modulation} = m \times 100\%. \qquad (3.3\text{–}7)$$

With $s_i(t)$ applied to the symmetric narrowband network with transfer function $H(j\omega)$, the output of the network $s_o(t)$ is given by

$$s_o(t) = [A(1 + m\cos\omega_m t) * h_L(t)]\cos\left[\omega_0 t + \theta(\omega_0)\right]. \qquad (3.3\text{–}8)$$

However,

$$A * h_L(t) = AH_L(0)$$

and

$$Am\cos\omega_m t * h_L(t) = Am|H_L(j\omega_m)|\cos\left[\omega_m t + \theta_L(\omega_m)\right],$$

where $H_L(j\omega)$ is the low-pass equivalent transfer function of $H(j\omega)$ and $\theta_L(\omega) = \arg H_L(j\omega)$. Therefore, $s_o(t)$ takes the form of a sinusoidally modulated AM signal given by

$$s_o(t) = AH_L(0)\left\{1 + \frac{m|H_L(j\omega_m)|}{H_L(0)}\cos\left[\omega_m t + \theta_L(\omega_m)\right]\right\}\cos\left[\omega_0 t + \theta(\omega_0)\right]. \qquad (3.3\text{–}9)$$

We note that the modulation index $m'(\omega_m)$ of $s_o(t)$ [obtained with the aid of Eq. (3.3–6)] is given by

$$m'(\omega_m) = m\frac{|H_L(j\omega)|}{H_L(0)}$$

and that the output modulation is shifted in phase by $\theta_L(\omega_m)$.

Example 3.3–1 For the circuit shown in Fig. 3.3–7, determine an expression for $v_o(t)$.

Solution. For the parallel RLC circuit,

$$\omega_0 = \frac{1}{\sqrt{LC}} = 10^7\ \text{rad/sec} \qquad \text{and} \qquad Q_T = \omega_0 RC = 10.$$

Since the circuit is symmetric about ω_0 we can obtain the envelope $v_{o_L}(t)$ of $v_o(t)$ by passing $(5\ \text{mA})(1 + \cos 5 \times 10^5 t)$ through the low-pass equivalent network, as shown in Fig. 3.3–8. The component of v_{o_L} due to the 5 mA constant input is simply

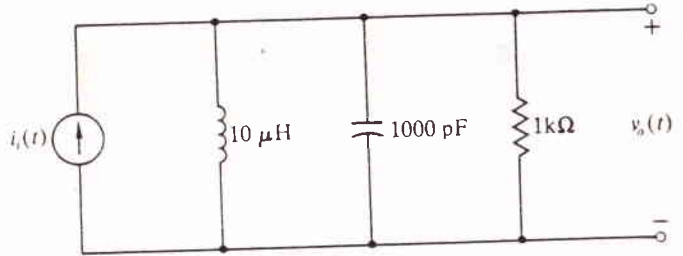


Figure 3.3–7

$i_i(t) = 5\text{mA}\,[1 + \cos 5 \times 10^5 t] \cos 10^7 t$

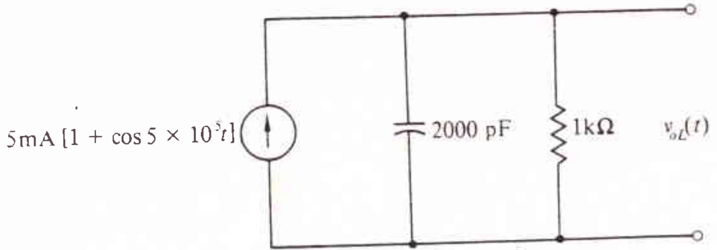


Figure 3.3–8

$\alpha = 5 \times 10^5$ rad/s

5 mA × 1 kΩ = 5 V, whereas the component of v_{oL} due to the (5 mA) cos 5 × $10^5 t$ of input is $[(5\text{ V})/\sqrt{2}] \cos(5 \times 10^5 t - \pi/4)$. The attenuation of $1/\sqrt{2}$ and the phase shift of $-\pi/4$ are both due to the fact that the input cosinusoid lies exactly at the −3 dB point of the low-pass filter. Combining the two components of $v_{oL}(t)$ and multiplying by $\cos \omega_0 t$, we finally obtain

$$v_o(t) = (5\text{ V})\left[1 + \frac{1}{\sqrt{2}} \cos\left(5 \times 10^5 t - \frac{\pi}{4}\right)\right] \cos 10^7 t.$$

This expression for $v_o(t)$ could also have been obtained by direct substitution into Eq. (3.3–9) of $A = 5$ mA, $H_L(0) = 1$ kΩ, $H_L(j\omega_m) = (1\text{ k}\Omega)/(1 + j) = (1/\sqrt{2})e^{-j(\pi/4)}$, and $\theta(\omega_0) = 0$.

Another example exploring the measurement of high-Q filters appears in the appendix at the end of the chapter.

3.4 NARROWBAND NETWORKS WITH PERIODIC INPUTS

One of the most efficient methods of amplifying a high-level sinusoidal signal is to convert the signal into a periodic train of narrow pulses in a nonlinear amplifier and then to pass these pulses through a narrowband filter to reconstruct the original sinusoid. In addition, one of the most fundamental methods of generating an amplitude-modulated wave is to control the amplitude of a periodic waveform (usually a square wave or a train of narrow pulses) and then to pass the wave through a narrowband filter to obtain a sinusoidal carrier. Since both of the above techniques

employ a narrowband filter with a periodic input, we shall determine the output $s_o(t)$ of such a filter centered at ω_0 when the input signal is of the form

$$s_i(t) = g(t)s_p(t),$$

where $s_p(t) = s_p(t + T)$ is periodic with period $T = 2\pi/\omega_0$ and $g(t)$ is the low-frequency signal which controls the envelope of $s_p(t)$. For high-level amplifier, $g(t)$ reduces to a constant.

Since $s_p(t)$ is periodic, it may be expanded in a Fourier series of the form

$$s_p(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t)$$

$$= a_0 + \sum_{n=1}^{\infty} C_n \cos(n\omega_0 t + \theta_n), \tag{3.4–1}$$

where $C_n = \sqrt{a_n^2 + b_n^2}$ and $\theta_n = -\tan^{-1}(b_n/a_n)$. The input signal to the filter therefore takes the form

$$s_i(t) = a_0 g(t) + \sum_{n=1}^{\infty} C_n g(t) \cos(n\omega_0 t + \theta_n), \tag{3.4–2}$$

which is an infinite superposition of AM waves, each centered at a harmonic ω_0. If the maximum frequency component ω_m of $g(t)$ is much less than ω_0 (which is almost true in practice), then the spectrum of each AM wave occupies a narrow band of frequencies of $2\omega_m$ about its center frequency as shown in Fig. 3.4–1. If $s_i(t)$ is now passed through a narrowband filter for which $|H(jn\omega_0)| \approx 0$ for $n = 0, 2, 3, 4, \ldots$, then the output $s_o(t)$ of the filter can be closely approximated by the response of the filter to only the fundamental ($n = 1$) component of $s_i(t)$; that is,

$$s_o(t) \approx [C_1 g(t) \cos(\omega_0 t + \theta_1)] * h(t), \tag{3.4–3}$$

where $h(t)$ is the impulse response of the narrowband filter. If the filter is symmetric about ω_0, then $s_o(t)$, with the aid of Eq. (3.3–3), can be written in the equivalent form

$$s_o(t) = [C_1 g(t) * h_L(t)] \cos[\omega_0 t + \theta_1 + \theta(\omega_0)], \tag{3.4–4}$$

where $h_L(t)$ is the impulse response of the low-pass equivalent filter and $\theta(\omega)$ is the phase angle of the narrowband filter. For the case where $g(t)$ is a constant or where

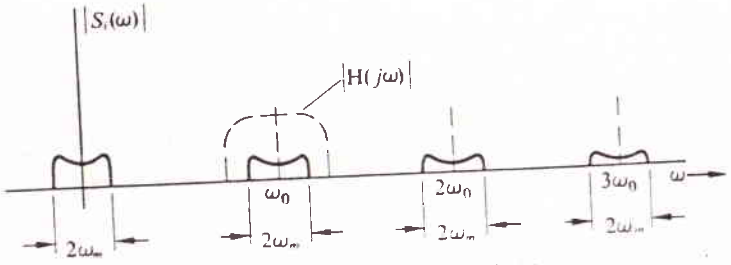


Fig. 3. 4–1 Typical spectrum of $s_i(t)$.

$_L(j\omega) \approx H_L(0)$ over the band of frequencies occupied by $g(t)$, [i.e., where the low-ass equivalent filter passes $g(t)$ undistorted], the filter output simplifies to

$$\begin{aligned} s_o(t) &= C_1 H_L(0) g(t) \cos[\omega_0 t + \theta_1 + \theta(\omega_0)] \\ &= C_1 |H(j\omega_0)| g(t) \cos[\omega_0 t + \theta_1 + \theta(\omega_0)], \end{aligned} \tag{3.4–5}$$

hich is in the expected form of an AM signal with envelope $g(t)$. The constant C_1, . of course, the fundamental component of the original periodic waveform and $H(j\omega_0)| = H_L(0)$ is the transfer function of the filter in the vicinity of the fundamental requency.

As a specific example of the above procedure, let us evaluate the output of the igh-Q parallel RLC circuit shown in Fig. 3.4–2, which is driven by a periodic train f current impulses applied at $t = 0$. The input current $i_i(t)$ has the form

$$i_i(t) = q \sum_{k=0}^{\infty} \delta(t - kT), \tag{3.4–6}$$

vhere q is the impulse strength (in coulombs) and $T = 2\pi/\omega_0$ is the spacing between mpulses. If we rewrite $i_i(t)$ in the equivalent form

$$i_i(t) = q u(t) s_p(t), \tag{3.4–7}$$

vhere $s_p(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)$, it becomes apparent that the input current is an nvelope-modulated periodic waveform.

To begin our analysis, we expand the periodic train of impulses $s_p(t)$ in a Fourier series to obtain

$$s_p(t) = \frac{1}{T} + \frac{2}{T} \sum_{n=1}^{\infty} \cos n\omega_0 t \tag{3.4–8}$$

and in turn

$$i_i(t) = I_0 u(t) + 2I_0 u(t) \sum_{n=1}^{\infty} \cos n\omega_0 t, \tag{3.4–9}$$

where $I_0 = q/T$ is the average value of $i_i(t)$ for $t > 0$. If

$$|Z_{11}(jn\omega_0)| \approx 0 \qquad \text{for} \qquad n = 0, 2, 3, 4, \ldots,$$

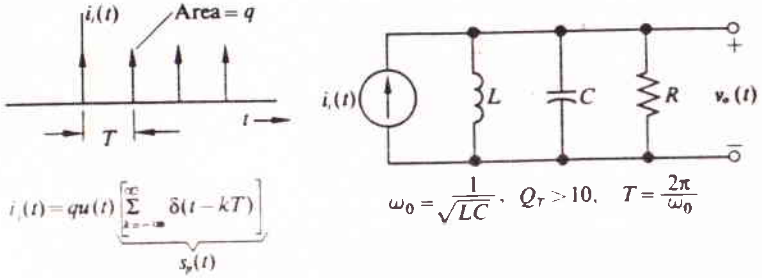

$i_i(t)$
Area = q
T
t
$i_i(t) = qu(t)\underbrace{\left[\sum_{k=-\infty}^{\infty} \delta(t-kT)\right]}_{s_p(t)}$
$i_i(t)$
L
C
R
+
$v_o(t)$
−
$\omega_0 = \frac{1}{\sqrt{LC}}$, $Q_T > 10$, $T = \frac{2\pi}{\omega_0}$

Fig. 3.4–2 Parallel RLC circuit driven by impulse train.

or equivalently

$$\frac{|Z_{11}(jn\omega_0)|}{|Z_{11}(j\omega_0)|} = \frac{|Z_{11}(jn\omega_0)|}{R} \ll 1 \qquad \text{for} \qquad n \neq 1,$$

then we need only retain the fundamental component of $i_i(t)$ to obtain an expression for $v_o(t)$. Physically, the components of $i_i(t)$ in the vicinity of $n\omega_0$ ($n \geq 2$) would be effectively shorted to ground through the capacitor C and thus would not contribute to $v_o(t)$, while the component of $i_i(t)$ in the vicinity of $\omega = 0$ would be effectively shorted to ground through the inductor L and thus would not contribute to $v_o(t)$. With the assumption that only the fundamental component of $i_i(t)$ contributes to $v_o(t)$ we may write

$$\begin{aligned} v_o(t) &\approx [2I_0 u(t) \cos \omega_0 t] * z_{11}(t) \\ &= [2I_0 u(t) * z_{11_L}(t)] \cos \omega_0 t \\ &= 2I_0 R(1 - e^{-\alpha t}) u(t) \cos \omega_0 t \qquad \text{[cf. Eq. (3.3–4)].} \end{aligned} \tag{3.3–10}$$

In practice, if one has a very narrow pulse in the train (so that all the harmonic terms have nearly equal amplitudes) and a broad filter (so that the response is not down too much at the harmonics), then the previous approach will be somewhat in error.† This error can be estimated from the size of the network transfer function at the various harmonics and the size of the input signal component at each harmonic.

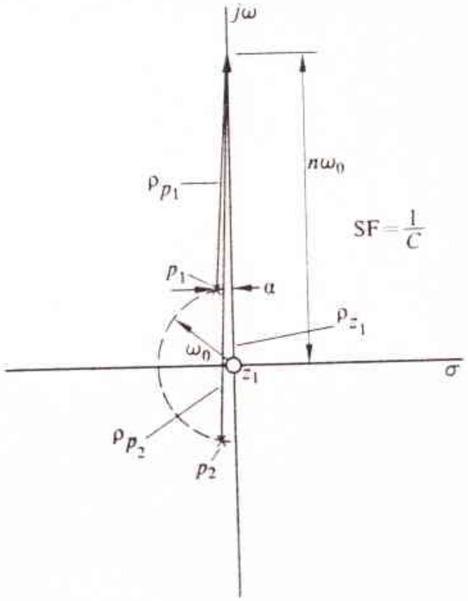


Fig. 3.4–3 Graphical evaluation of $|Z_{11}(j\omega)|$.

---

† From the time viewpoint, the addition of an impulse to a network is bound to cause a stepwise, rather than a smooth, increment in the stored energy and hence in the waveshape appearing across the network.

s an example, it is relatively simple to evaluate $|Z_{11}(jn\omega)|$ for any narrowband rk from a sketch of its pole-zero diagram. Figure 3.4–3 shows the case for a el RLC circuit: $|Z_{11}(jn\omega)|$ is found as a scale factor ($1/C$ in this case) times the tude of the phasor from the zero to $n\omega$ on the imaginary axis divided by the ct of the magnitudes of the phasors from the two complex poles to the same If $Q_T > 10$, or equivalently $\alpha \ll \omega_0$, we observe that

$$\rho_{p_1} \approx (n-1)\omega_0, \qquad \rho_{p_2} \approx (n+1)\omega_0, \qquad \text{and} \qquad \rho_{z_1} = n\omega_0;$$

$$|Z_{11}(jn\omega_0)| \approx \frac{(1/C)\rho_{z_1}}{\rho_{p_1}\rho_{p_2}} = \frac{n}{\omega_0 C(n^2-1)}. \tag{3.4–11}$$

ows that

$$\frac{|Z_{11}(jn\omega_0)|}{|Z_{11}(j\omega_0)|} = \frac{n}{\omega_0 RC(n^2-1)} = \frac{n}{Q_T(n^2-1)}. \tag{3.4–12}$$

e might expect, as Q_T increases, the ratio of $|Z_{11}(jn\omega_0)|$ to $|Z_{11}(j\omega_0)|$ decreases. fically for $Q_T \geq 10$,

$$\left|\frac{Z_{11}(jn\omega_0)}{Z_{11}(j\omega_0)}\right| < 0.067 \qquad \text{for all} \qquad n \neq 1.$$

Using these results one can estimate the minimum Q necessary to keep the rms ibution of the second- and third-harmonic terms of an ideal impulse train ied to a parallel tuned circuit below 1% of the fundamental contribution to the ut. Since all components of an impulse pulse train are equal in magnitude, the ed value of Q is a solution of the equation

$$\frac{2}{3Q_T}\sqrt{1 + (\tfrac{3}{8} \times \tfrac{3}{2})^2} \leq \tfrac{1}{100},$$

$$Q_T \geq \sqrt{\tfrac{337}{576}}\,100 \approx 76.$$

Since all other waveshapes have harmonics that fall off in magnitude as the uency increases, it follows that any waveshape other than an impulse in the train not require as high a value of Q to keep the distortion from the second and third monics down to 1%. The distortion problem is considered again in a more nal and detailed manner in Section 3.5.

TOTAL HARMONIC DISTORTION

periodic signal $s_p(t)$ of period T drives a narrowband filter centered at $\omega_0 = 2\pi/T$. filter essentially extracts the fundamental component of $s_p(t)$. However, if the er transfer function is not identically equal to zero at the harmonics of ω_0, some

harmonic components appear at the filter output. Specifically, if

$$s_p(t) = C_0 + \sum_{n=1}^{\infty} C_n \cos(n\omega_0 t + \theta_n) \tag{3.5–1}$$

and if $H(j\omega)$ is the transfer function of the narrowband filter, then the filter output $s_o(t)$ has the form

$$s_o(t) = C_0 H(0) + \sum_{n=1}^{\infty} C_n |H(jn\omega_0)| \cos[n\omega_0 t + \theta_n + \theta(n\omega_0)], \tag{3.5–2}$$

where $\theta(n\omega_0) = \arg H(jn\omega_0)$. In many applications, in particular that of obtaining a pure sinusoid at the output of an oscillator, the higher harmonics of $s_o(t)$ represent an unwanted distortion signal $s_d(t)$, where $s_d(t)$ may be written in the form

$$s_d(t) = \sum_{n=2}^{\infty} C_n |H(jn\omega_0)| \cos[n\omega_0 t + \theta_n + \theta(n\omega_0)]. \tag{3.5–3}$$

[The $C_0 H(0)$ term has been deliberately omitted from $s_d(t)$, since it can always be removed by an RF choke or a blocking capacitor.]

To obtain a quantitative measure of the effectiveness of a narrowband filter in reducing $s_d(t)$ relative to the fundamental component of $s_o(t)$, we define *total harmonic distortion* (THD) as the ratio of the rms value of $s_d(t)$ to the fundamental component of $s_o(t)$. Thus

$$\text{THD} = \frac{(s_d)_{\text{rms}}}{(s_o - s_d - \bar{s}_o)_{\text{rms}}} = \sqrt{\sum_{n=2}^{\infty} \left(\frac{C_n}{C_1}\right)^2 \left[\frac{|H(jn\omega_0)|}{|H(j\omega_0)|}\right]^2}, \tag{3.5–4}$$

where $\sqrt{\sum_{n=2}^{\infty} (C_n^2/2)|H(jn\omega_0)|^2}$ is the rms value of $s_d(t)$. The quantity THD given by the above equation is exactly the meter reading that would appear on a distortion analyzer driven by $s_o(t)$. It is apparent that for a given input signal $s_p(t)$, the smaller the value of THD the better the filter performs in producing a pure sinusoidal output.

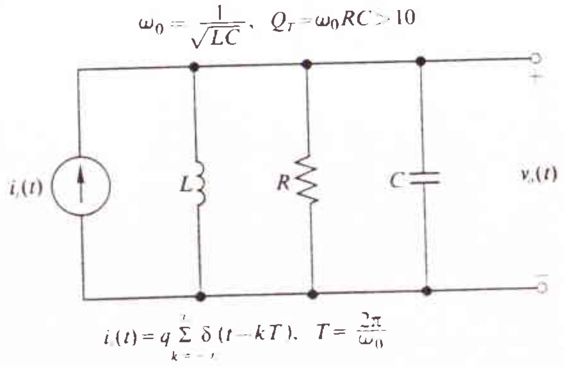

$\omega_0 = \frac{1}{\sqrt{LC}}, \quad Q_T = \omega_0 RC > 10$
$i_i(t)$
L
R
C
$v_o(t)$
+
−
$i_i(t) = q \sum_{k=-\infty}^{\infty} \delta(t-kT), \quad T = \frac{2\pi}{\omega_0}$

Fig. 3.5–1 High-Q parallel RLC circuit driven by a periodic train of impulses.

To illustrate the method of obtaining an expression for THD, let us examine the tput $v_o(t)$ of the high-Q_T parallel RLC circuit shown in Fig. 3.5–1. For this circuit periodic input current may be expanded in a Fourier series of the form

$$i_i(t) = q \sum_{k=-\infty}^{\infty} \delta(t - kT) = C_0 + \sum_{n=1}^{\infty} C_n \cos n\omega_0 t,$$

iere $C_n = 2C_0 = 2q/T$; hence $C_n/C_1 = 1$ for $n = 2, 3, \ldots$. In addition, for a gh-Q_T parallel RLC circuit,

$$\frac{|Z_{11}(jn\omega_0)|}{|Z_{11}(j\omega_0)|} = \frac{n}{Q_T(n^2 - 1)}$$

. Eq. 3.4–12); thus

$$\text{THD} = \frac{1}{Q_T}\sqrt{\sum_{n=2}^{\infty}\left(\frac{n}{n^2 - 1}\right)^2}. \tag{3.5–5}$$

he series $\sum_{n=2}^{\infty} (n/n^2 - 1)^2$ is known to converge to $(1/16 + \pi^2/12)$,† which when ıbstituted into Eq. (3.5–5), yields

$$\text{THD} = \frac{1}{Q_T}\sqrt{\frac{1}{16} + \frac{\pi^2}{12}} \approx \frac{0.94}{Q_T}. \tag{3.5–6}$$

Ve now observe that for THD less than 1%, Q_T must be greater than 94. (A value f 76 was obtained in the previous section by considering only the second- and third-ıarmonic distortion terms.)

If in lieu of a periodic train of impulses, $i_i(t)$ in the circuit of Fig. 3.5–1 is a square vave with a peak-to-peak amplitude I, a period $T = 2\pi/\omega_0$, and zero average value, hen

$$i_i(t) = \sum_{n=1}^{\infty} C_{2n-1} \cos (2n - 1)\omega_0 t, \tag{3.5–7}$$

vhere $C_{2n-1} = 2I(-1)^{n-1}/\pi(2n - 1)$. For this case $(C_{2n-1}/C_1)^2 = (1/2n - 1)^2$; thus

$$\begin{aligned}\text{THD} &= \frac{1}{Q_T}\sqrt{\sum_{n=2}^{\infty}\left[\frac{2n-1}{(2n-1)^2 - 1}\right]^2\left[\frac{1}{2n-1}\right]^2} \\ &= \frac{1}{Q_T}\sqrt{\sum_{n=2}^{\infty}\left[\frac{1}{(2n-1)^2 - 1}\right]^2}.\end{aligned} \tag{3.5–8}$$

The series

$$\sum_{n=2}^{\infty}\left[\frac{1}{(2n-1)^2 - 1}\right]^2$$

---

† Although the series in Eq. (3.5–5) has been obtained in closed form for this case, such a representaion is in general not possible. Fortunately, in most practical cases, the series of Eq. (3.5–4) converges sufficiently rapidly so that the first two or three terms provide a reasonably good approximation to the entire series.

---

is known to converge to $(\pi^2/48 - 3/16)$; hence Eq. (3.5–8) simplifies to

$$\text{THD} = \frac{1}{Q_T}\sqrt{\frac{\pi^2}{48} - \frac{3}{16}} \approx \frac{0.135}{Q_T}. \tag{3.5–9}$$

To yield less than 1% THD with a square-wave drive to a parallel RLC circuit, Q_T must only exceed 13.5, which is a factor of 7 lower than the value of Q_T required to yield less than 1% THD with an impulse train drive. This reduction in the required value of Q_T results both from the fact that the even harmonic components of a square wave are zero and from the fact that the higher harmonics of a square wave have amplitudes that are small compared with the fundamental amplitude. (The amplitudes of the harmonics vary inversely with frequency.) Since a periodic train of impulses is the only waveform whose harmonic amplitudes do not decrease with increasing harmonic number, we expect the THD obtained with an impulse train drive to a particular narrowband filter to provide an upper bound on the THD obtained with any other periodic input drive; hence $Q_T = 94$ ensures less than 1% THD for a parallel RLC circuit regardless of the form of the periodic input current drive.

PROBLEMS

3.1 Find the low-pass equivalent circuit, from the $Y_{11}(p)$ viewpoint, for a high-Q ($Q_T \geq 10$) series RLC circuit.

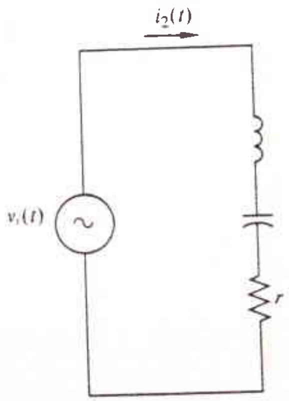


Figure 3.P–1

3.2 Determine $i_2(t)$ in Fig. 3.P–1, assuming that $\omega_0 = 10^7$ rad/sec, $r = 50\ \Omega$, $Q_T = 20$, and $v_i(t)$ is a 5 V peak cosine signal at $\omega_0 = 10^7$ rad/sec that is 100% amplitude modulated by a radian frequency of 2.5×10^5 rad/sec.

3.3 Determine the output voltages $v_o(t)$ when a unit impulse, a unit step, and a unit step modulated sine wave of $f_o = 1.6$ MHz are applied successively as $i(t)$ in the circuit of Fig. 3.P–2.

3.4 Find the low-pass equivalent circuit of the high-Q circuit shown in Fig. 3.P–3.

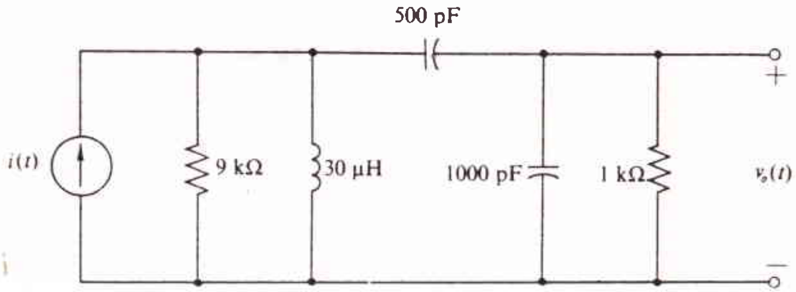


Figure 3.P–2

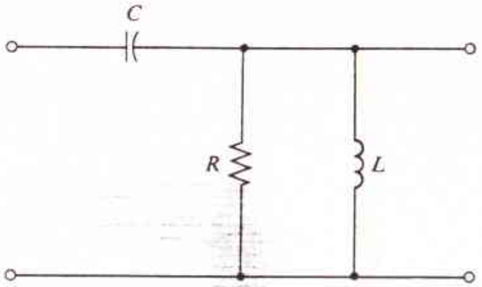
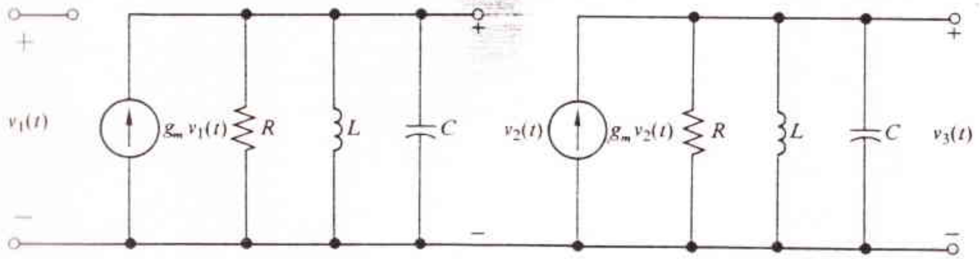


Figure 3.P–3



$\omega_0 = \frac{1}{\sqrt{LC}}$ $Q_T = \omega_0 RC > 10$

Figure 3.P–4

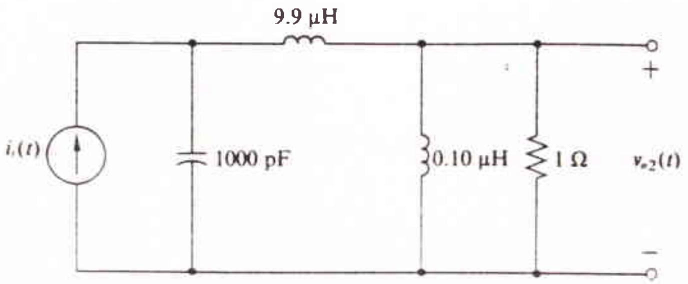


Figure 3.P–5

3.5 Find the response of the two-stage IF strip shown in Fig. 3.P–4 to an amplitude-modulated sine wave at the resonant frequency of the tuned circuits. Assume an input modulation of 100% and that the carrier is centered in the passband of the filter. Find the output percentage modulation for radian modulation frequencies of $1/4RC$, $1/2RC$, and $1/RC$. (Use the low-pass equivalent approach.)

3.6 A train of rectangular narrow pulses 0.1 μsec wide, 10 mA in amplitude, and spaced 628 nsec apart is applied as $i_i(t)$ to the circuit of Fig. 3.P–5. Estimate accurately the component of $v_{o2}(t)$ at 2×10^7 rad/sec. (*Hint*: The curve for $I_2/I_{\max}$ for a rectangular pulse in the Appendix to Chapter 9 may prove helpful if the pulse train problem is unfamiliar.)

3.7 Determine $v_o(t)$ for the circuit of Fig. 3.P–6 when $i_i(t) = (1\text{ mA})(\cos 2.5 \times 10^5 t) \cos 10^7 t$.

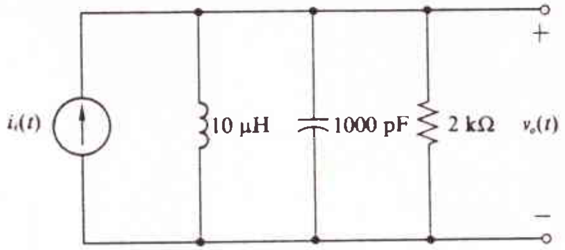


Figure 3.P–6

APPENDIX TO CHAPTER 3

HIGH-Q FILTER MEASUREMENTS

The results of Section 3.3 find practical application in the measurement of the magnitude and phase angle, as a function of ω_m, of the transfer function of extremely narrowband filters, i.e., filters having Q as high as 10^4. The frequency of a normal oscillator often cannot be adjusted with the required precision to make such measurements; and even if it could be so adjusted, it would be unlikely to possess the required frequency stability to allow the measurement to be made conveniently.

On the other hand, a crystal oscillator with the same center frequency as the filter of interest can possess the required frequency stability. If this oscillator is amplitude-modulated with the output of a low-frequency constant-amplitude variable-frequency oscillator, the resultant signal of the form

$$s_i(t) = A(1 + \cos \omega_m t) \cos \omega_0 t$$

does provide the test signal with which to evaluate the high-Q filter. With $s_i(t)$ applied to the filter, the output signal takes the form (Eq. 3.3–9)

$$s_o(t) = AH_L(0)\left\{1 + \underbrace{\frac{|H_L(j\omega_m)|}{H_L(0)}}_{m'(\omega_m)} \cos\left[\omega_m t + \theta_L(\omega_m)\right]\right\} \cos\left[\omega_0 t + \theta(\omega_0)\right]. \qquad (3.\text{A}–1)$$

If $m'(\omega_m) = |H_L(j\omega_m)|/H_L(0)$ and $\theta_L(\omega_m)$ are measured as a function of ω_m and if $H_L(0)$ and $\theta(\omega_0)$ are determined, then sufficient data exist to plot the magnitude and phase of $H(j\omega)$. Specifically, for $\omega > 0$,

$$|H(j\omega)| = |H_L(j\omega - \omega_0)|,$$

where $|H_L(j\omega_m)| = m'(\omega_m)H_L(0)$ and $\theta(\omega) = \theta_L(\omega - \omega_0) + \theta(\omega_0)$.

Experimentally $H_L(0)$ and $\theta(\omega_0)$ can be readily determined by any number of standard methods with the modulation on the crystal oscillator absent. In addition, $m'(\omega_m)$ and $\theta_L(\omega_m)$ can be determined by forming a Lissajous pattern on an oscilloscope face with $s_o(t)$ applied to the vertical channel and $\cos \omega_m t$ applied to the horizontal channel. A typical Lissajous pattern is shown in Fig. 3.A–1. Since the upper and lower envelopes of $s_o(t)$ form standard elliptical Lissajous patterns, $\theta_L(\omega_m)$ may be found from

$$\theta_L(\omega_m) = \sin^{-1}\frac{F}{E}. \qquad (3.\text{A}–2)$$

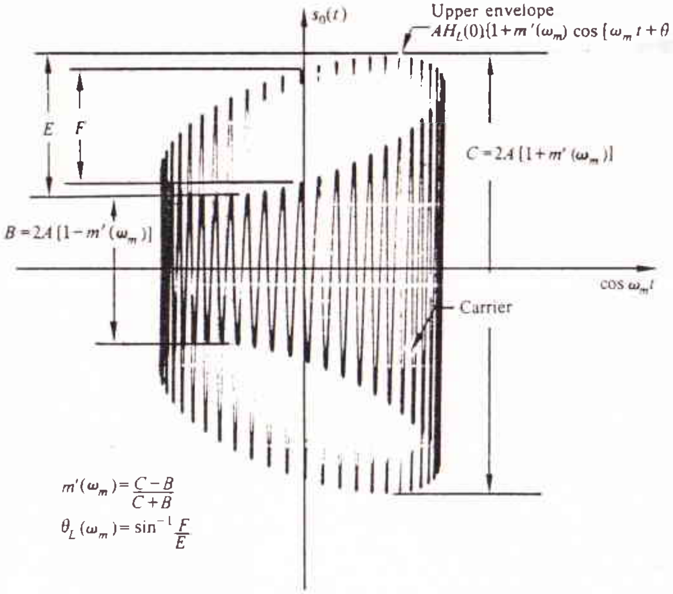

$s_0(t)$
Upper envelope
$AH_L(0)\{1+m'(\omega_m) \cos[\omega_m t+\theta(\omega_m)]\}$
E
F
$C = 2A[1+m'(\omega_m)]$
$B = 2A[1-m'(\omega_m)]$
$\cos \omega_m t$
Carrier
$m'(\omega_m) = \frac{C-B}{C+B}$
$\theta_L(\omega_m) = \sin^{-1}\frac{F}{E}$

Fig. 3.A–1 Lissajous pattern of $s_o(t)$ vs. $\cos \omega_m t$.

where E and F are indicated in Fig. 3.A–1. In addition, $m'(\omega_m)$ can be obtained as

$$m'(\omega_m) = \frac{C - B}{C + B} \qquad (3.\text{A}–3)$$

by employing Eq. (3.3–6). [The reader should convince himself of the validity of Eq. (3.A–2).] Consequently, by selecting a sufficient number of different values for ω_m, one can plot the magnitude and phase angle of $H(j\omega)$ to any degree of accuracy desired.

CHAPTER 4

NONLINEAR CONTROLLED SOURCES

The normal mode of operation of the active devices discussed in this book is nonlinear. Thus the usual small-signal incremental model for a device is not likely to be of much use either in designing or in analyzing a circuit. The purpose of this chapter is to remedy this situation by providing adequate large-signal models for a number of useful devices. We shall accomplish this aim by exploring general methods of dealing with the nonlinearities that occur in a number of "real-life" active devices or circuits. In order to concentrate on basic concepts, we assume the absence of device reactance; e.g., no charge storage is allowed in transistors. We also assume that only a single type of nonlinearity occurs in any given case. In the early sections, all sources are assumed to be ideal in that series or parallel losses are excluded. In later sections we examine the modifications that occur through the inherent or deliberate addition of resistive terms.

Chapter 5 will consider the combination of nonlinear and reactive effects, while all the following chapters will be concerned with complete circuits or systems.

4.1 GENERAL COMMENTS

One can break down the types of nonlinearities into two broad classes. In one class, the input-output relationship is of the piecewise-linear or switched-linear segment form; in the other, the relationship varies gradually and lacks abrupt changes of slope. Piecewise-linear types are always easily expressible in analytic form, as are certain of the gradual relationships. In some cases one can view a particular piecewise configuration as an asymptotic limit for a gradual expression.

For some device characteristics, simple analytic means of characterizing their relationships do not seem to exist. However, in these cases one can always, at least in theory, fit a polynomial to any desired degree over any desired range. Although computer programs to aid in such curve fitting are widely available, they should have to be resorted to only on rare occasions.

In general, one cannot resort to superposition in dealing with nonlinear circuits. This means that the dc (average-value) and ac (time-variable) components of the output signal are interrelated. Initially we deal with this problem by assuming ideal voltage or current sources as bias supplies, so that average values may be fixed independently of the time-variable terms. The circuit of Fig. 1.1-1 provides an example of such a bias circuit.

In addition to the ac-dc interrelationships of nonlinear sources, there are input drive waveshape–output waveshape relationships that normally defy generalization. This problem is circumvented in practice by assuming one of several common driving waveshapes and determining the outputs for them. Fortunately, it turns out that dc plus a sine-wave drive or dc plus a square-wave drive suffices to provide a first-order approximation for the operation of almost all real circuits.

4.2 PIECEWISE-LINEAR SOURCE, SINGLE DISCONTINUITY

Consider the circuit of Fig. 4.2–1. When $v_1 > V_0$, $i_2 = G(v_1 - V_0)$; when $v_1 \leq V_0$, $i_2 = 0$; hence, if

$$v_1 = V_b + v(t) \qquad \text{and} \qquad V_b - |v(t)|_{\max} > V_0,$$

then the operation occurs completely along the sloping portion of the characteristic and is incrementally linear in the conventional sense. In this case, superposition applies to the ac signal components; that is, if $i_2 = I_b + i(t)$, then $i(t) = Gv(t)$.

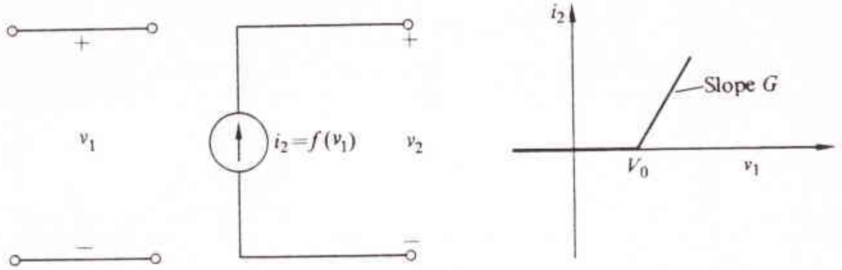


Fig. 4.2–1 Piecewise-linear voltage-controlled current source.

As another special case, consider the previous circuit biased exactly at the break point, so that $V_b = V_0$. Under this circumstance, the circuit performs as an ideal half-wave rectifier. In this case, if $v(t) = V_1 \cos \omega t$, then the output current consists of half-cycle sine-wave pulses of peak amplitude $I_p = GV_1$, while a square-wave input for $v(t)$ of peak-to-peak amplitude $2V_1$ and period $T = 2\pi/\omega$ yields a square-wave output also of peak value $I_p = GV_1$. Note that in the sine-wave case the output waveshape is drastically different from the input waveshape over one half-cycle, whereas it has the identical shape over the other half-cycle. On the other hand, a square-wave input yields a square-wave output.

In many practical circuits the drive is periodic and the nonlinear device is followed by a narrow band filter of the type discussed in Chapters 2 and 3. In these cases it is convenient to express the device output in a Fourier-series form that clearly displays the various frequency components and their phase relationships.

This is very easy to do when $V_b = V_0$ and $v(t) = V_1 \cos \omega t$, since the Fourier-series expansion of a series of half-sine-wave pulses is given by

$$i_2(t) = \frac{I_p}{\pi} + \frac{I_p}{2} \cos \omega t + \frac{2I_p}{3\pi} \cos 2\omega t - \frac{2I_p}{15\pi} \cos 4\omega t + \cdots, \tag{4.2-1}$$

where $I_p = GV_1$ is the peak value of i_2. With $v(t)$ in the form of a square wave,

$$i_2(t) = \frac{I_p}{2} + \frac{2I_p}{\pi} \cos \omega t - \frac{2I_p}{3\pi} \cos 3\omega t + \frac{2I_p}{5\pi} \cos 5\omega t - \cdots, \tag{4.2-2}$$

where again $I_p = GV_1$ and the time origin is taken at the middle of the positive pulse.

It is worth noting at this point that if the current generator of Fig. 4.2–1 drives a parallel resonant circuit that is tuned to ω and that has a parallel resistance R_T and a high value of Q_T, then only the fundamental term produces any significant voltage v_2 across the tuned circuit. Consequently, with $V_b = V_0$, $v_2(t)$ is given by

$$\text{sine-wave drive:} \qquad v_2 \approx \left(\frac{V_1 G R_T}{2}\right) \cos \omega t, \tag{4.2-3}$$

$$\text{square-wave drive:} \qquad v_2 \approx \left(\frac{2V_1 G R_T}{\pi}\right) \cos \omega t. \tag{4.2-4}$$

With both the square-wave and sine-wave drive we have obtained a sinusoidal output voltage that is *linearly* related in amplitude to the input driving voltage amplitude V_1. Thus a highly nonlinear device operation has been combined with a narrowband filter to produce an overall linear amplifier. Such devices find wide application in the efficient amplification of AM waves. The "gain" of this amplifier is a function of the driving waveshape.

As we shall see in Chapter 9, amplifiers in which the output current flows for exactly half of each input cycle are known as Class B amplifiers. From the equations above, we see that a Class B RF amplifier that has a linear characteristic over its conducting half-cycle has an overall linear output–input characteristic in spite of its highly nonlinear internal behavior. This is but the first of many examples in which we discover overall "linearity" in spite of internal nonlinearities.

We now consider the general case of sine- or square-wave drives added to an arbitrary dc bias V_b. The square-wave case is trivial, since it always results in a square-wave output and hence Eq. (4.2–2) is always valid. All that changes from case to case is the peak amplitude of the output square wave. For example, if $V_b \neq V_0$, then $I_p = G(V_b + V_1 - V_0)$.

In the sine-wave case the waveshape is a function of the bias V_b and the amplitude V_1; hence Eq. (4.2–1) is not valid for any case except the half-wave rectifier. The interrelationship of i_2, V_b, V_0, and V_1 is shown in Fig. 4.2–2. Here we see that i_2 is in the form of a periodic train of sine-wave tips of peak amplitude $I_p = (V_1 - V_x)G$, where $V_x = V_0 - V_b$. In addition, if i_2 is plotted vs. ωt rather than t, then an entire cycle occupies a duration of 2π and the sine-wave tip occupies a duration of 2ϕ, which we define as the conduction angle, where

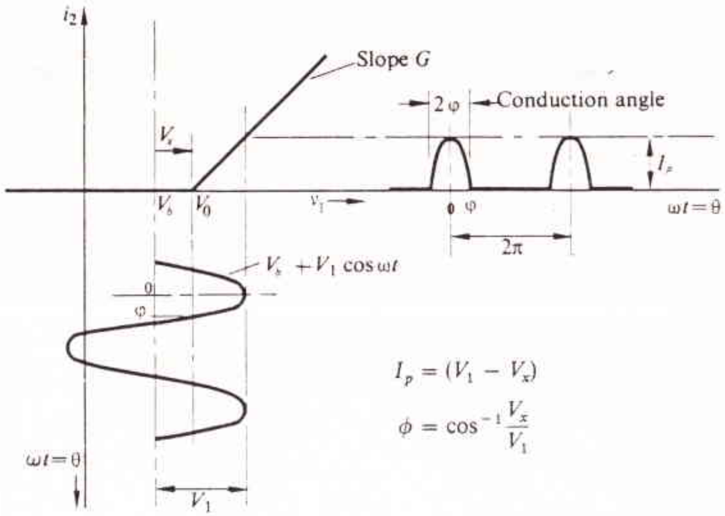


Fig. 4.2–2 Output waveshape for the circuit of Fig. 4.2–1.

$$\phi = \cos^{-1} \frac{V_x}{V_1}.$$

We observe that for $V_b \geq V_0 + V_1$, or equivalently $V_x \leq -V_1$, the conduction angle occupies an entire cycle and incrementally linear operation occurs.

If we write a general expression for the Fourier series expansion of a train of sine-wave tips in the form†

$$i_2 = \sum_{n=0}^{\infty} I_n \cos n\omega t,$$

then this expansion may be used to determine the dc, fundamental, or harmonic content in the output of a piecewise-linear circuit driven by dc plus a sine wave.

The algebraic expression for the coefficients I_n of such an expansion and their asymptotic values are presented in the appendix to this chapter. Figure 4.2–3 presents normalized values for these coefficients in terms of the conduction angle 2ϕ and the peak pulse amplitude I_p. The plus or minus sign after the coefficient number indicates whether the coefficient is positive or negative.

As a practical matter, it is often more convenient to have the Fourier series coefficients expressed in terms of V_x/V_1 than in terms of ϕ, since in most problems ϕ is not known explicitly. This transposition of coordinate axes is readily

---

† We should recall that a cosinusoid applied to a nonlinear nonmemory device produces a periodic output which may be expanded in a Fourier *cosine* series with no sine terms.

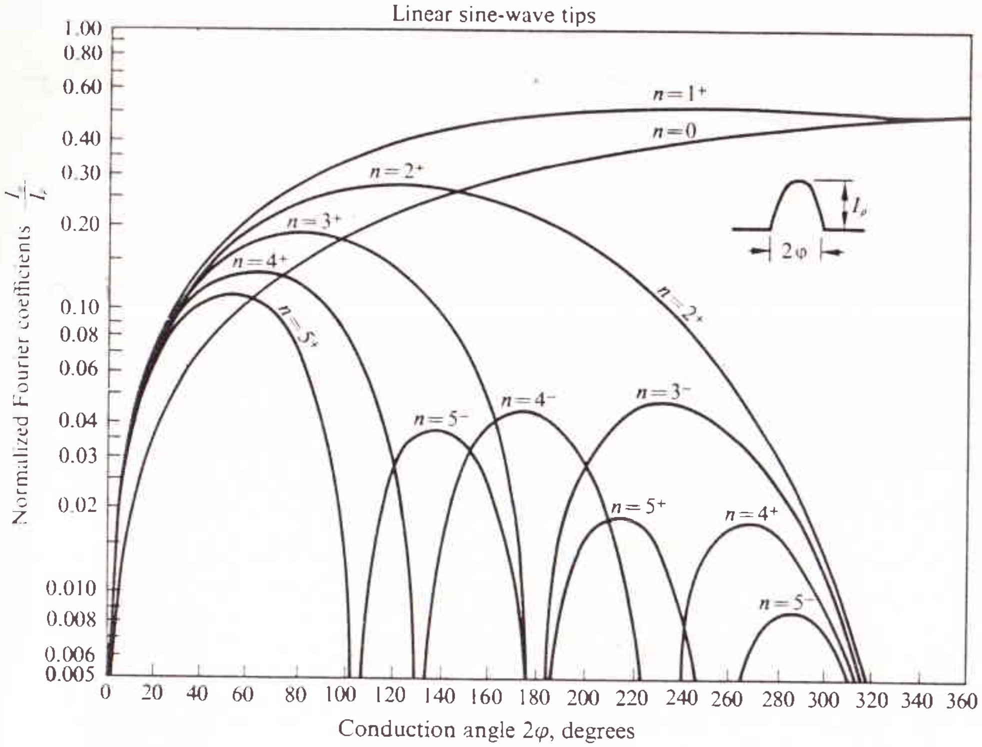


Fig. 4.2–3 Normalized Fourier coefficients of a linear sine-wave tip pulse train vs. the conduction angle.

accomplished by employing the relationship $\cos\phi = V_x/V_1$. In particular, I_n/I_p for $n = 0, 1, 2$ is plotted vs. V_x/V_1 in Fig. 4.2–4. Negative values of V_x/V_1 in this figure correspond to the case where $V_b > V_0$ (note that $V_x = V_0 - V_b$) and i_2 has a conduction angle greater than 180°.

As an example of the usefulness of these curves, consider a case in which $V_x = 1.5$ V, $G = 1000\ \mu$mho, and $V_1 = 4.5$ V. From the data of Fig. 4.2–4 and by multiplying by $I_p = G(V_1 - V_x) = 3$ mA, we obtain $I_0 = 0.78$ mA, $I_1 = 1.32$ mA, and $I_2 = 0.81$ mA. If this current were passed through a tuned circuit or its equivalent, then the various output voltages could be found by performing one more simple multiplication for each current component.

4.3 MULTIPLE-SEGMENT PIECEWISE-LINEAR SOURCES

With large enough input drives, all practical devices eventually "saturate." In many cases the saturation is sufficiently abrupt that the input-output transfer characteristic, v_2 vs. i_1, can be modeled by a multiple-segment piecewise-linear characteristic of the type shown in Fig. 4.3–1.

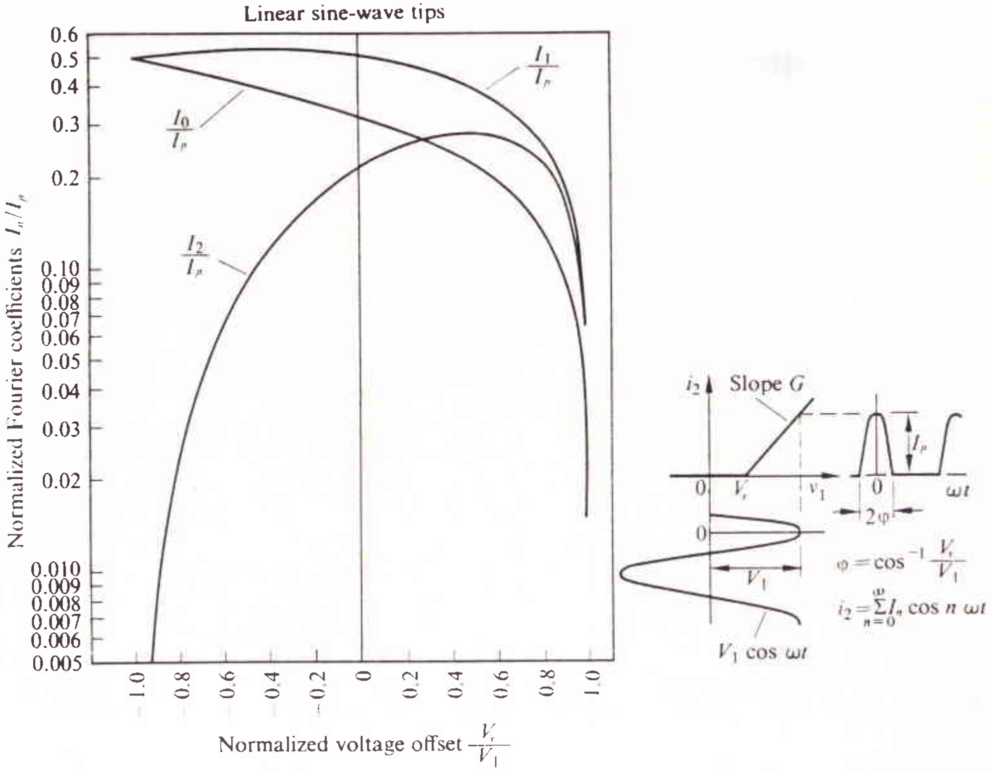


Fig. 4.2–4 Plot of I_n/I_p vs. V_x/V_1 for $n = 0, 1,$ and 2.

Results for the case shown in Fig. 4.3–1(a) can be found as the superposition of the Fourier series for the positive and the negative sine-wave tip pulse trains. For the positive pulse train i_{2P} of peak amplitude $I_{p1} = G_1(V_1 - V_x)$, we obtain the Fourier coefficients from Fig. 4.2–4 by employing I_{p_1}, and V_x/V_1. The Fourier coefficients for the negative pulse train i_{2N}, of peak amplitude $I_{p_2} = G_2(V_1 + V_y)$, are obtained by employing I_{p_2} and $-V_y/V_1$ (in lieu of V_x/V_1) in the curves of Fig. 4.2–4. The negative pulse train is shifted 180° in phase; therefore, its Fourier series has the form

$$\begin{aligned} i_{2N} &= -[I_0 + I_1\cos(\omega t + \pi) + I_2\cos(2\omega t + 2\pi) + \cdots,] \\ &= -I_0 + I_1\cos\omega t - I_2\cos 2\omega t + I_3\cos 3\omega t - \cdots. \qquad (4.3\text{–}1) \end{aligned}$$

Consequently, when i_{2N} is combined with i_{2P}, the Fourier coefficients of the odd harmonics of both pulse trains add algebraically, whereas the Fourier coefficients of the even harmonics and the average value of the negative pulse train must be subtracted from the corresponding coefficients of the positive pulse train to obtain the composite Fourier coefficients.

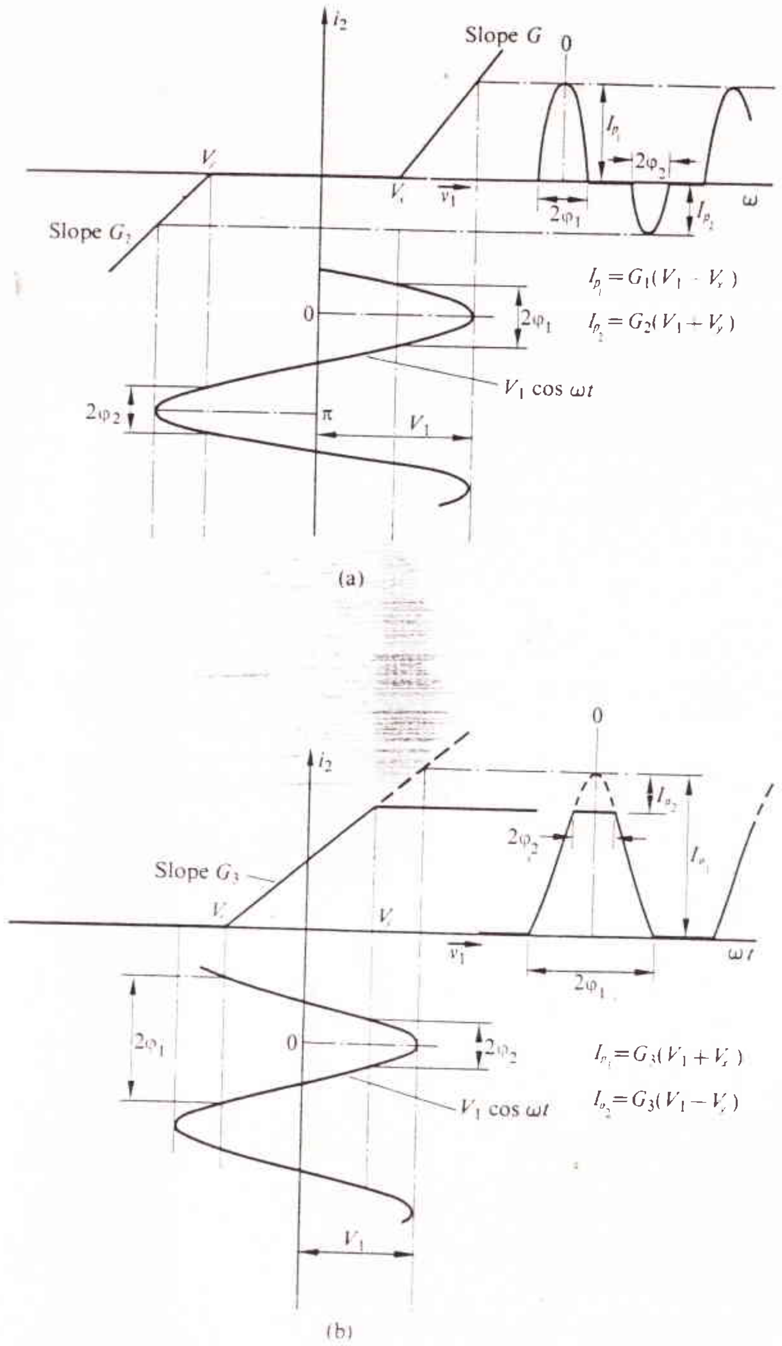


Fig. 4.3–1 Various multiple-segment piecewise-linear nonlinearities, dc plus sine-wave drive.

We should note that if $G_1 = G_2$ and $V_x = -V_y$, no average value or even harmonics exist in the waveform of i_2, since the magnitude of the corresponding Fourier coefficients of both the positive and negative pulse trains are equal.

Results for the saturation model shown in Fig. 4.3–1(b) can be found by direct subtraction of the Fourier series for a sine-wave tip pulse train of peak amplitude $I_{p_2} = G_3(V_1 - V_y)$ and conduction angle $2\phi_2[\phi_2 = \cos^{-1}(V_y/V_1)]$ from the Fourier series for a sine-wave tip pulse train of peak amplitude $I_{p_1} = G_3(V_1 - V_x)$ and the conduction angle $2\phi_1[\phi_1 = \cos^{-1}(V_x/V_1)]$. Subtraction of the two Fourier series results in a Fourier series whose coefficients are found from the subtraction of the corresponding coefficients of the original two series. The coefficients of each series are, of course, found directly from Fig. 4.2–3 or Fig. 4.2–4. In the limit as V_1 becomes very large the current waveshape approaches a square wave; hence Eq. (4.2–2) again yields limiting values.

Suppose, for example, that $V_y = -V_x = 1.5$ V, that the sine wave is varying around zero with an amplitude $V_1 = 3$ V, and that $G_3 = 500\,\mu$mho. Then

$$\frac{V_x}{V_1} = \frac{-V_y}{V_1} = -0.5,$$

$$I_{p_1} = [500 \times 10^{-6} \times 4.5]\text{ mA} = 2.25\text{ mA},$$

and

$$I_{p_2} = [500 \times 10^{-6} \times 1.5]\text{ mA} = 0.75\text{ mA}.$$

With the aid of Fig. 4.2–4 we obtain

$$I_0 = (2.25\text{ mA}) \times 0.41 - (0.75\text{ mA}) \times 0.226 = 0.75\text{ mA}$$

[which, of course, could be found by inspection from Fig. 4.3–1(b)],

$$I_1 = (2.25\text{ mA}) \times 0.535 - (0.75\text{ mA}) \times 0.40 = 0.91\text{ mA},$$

and

$$I_2 = (2.25\text{ mA}) \times 0.093 - (0.75\text{ mA}) \times 0.28 = 0.00\text{ mA}$$

(again as expected because of the symmetry).

In general, the output of any multiple-segment piecewise-linear characteristic driven by dc plus a sine wave may be expressed as the superposition of several sine-wave tip pulse trains. A little ingenuity is sometimes required to accomplish this. Once this representation is achieved, however, the coefficients of the output Fourier series may be obtained with the aid of the curves of Fig. 4.2–3 or Fig. 4.2–4.

Characteristics similar to Fig. 4.3–1(a) occur in Class B and Class C amplifiers, whereas center-biased characteristics similar to Fig. 4.3–1(b) are used as limiters. If a characteristic like that of Fig. 4.3–1(b) is really biased in the middle, then there are no even harmonics at the output. In addition, the normalized fundamental current output I_1/G_3V_y varies only from 1.0 to $4/\pi = 1.273$ (the fundamental of a square wave of peak amplitude = 1) as the input amplitude varies from the break-point value, $V_y = -V_x$, to infinity. For $V_1 = 2V_y$,

$$I_{p_1} = 3G_3V_y, \qquad I_{p_2} = G_3V_y, \qquad \frac{V_y}{V_1} = 0.5, \qquad \frac{V_x}{V_1} = -0.5,$$

and thus

$$G_3V_y/I_1 = 1.218.$$

Similarly, if $V_1 = 3.86V_y$, then the normalized current is 1.258. A plot of I_1/G_3V_y vs. V_1/V_y is shown in Fig. 4.3–2. It is apparent that if an amplitude-modulated signal is applied to such a characteristic so that the carrier amplitude is 10 times the break-point while the percentage amplitude modulation is 80% or less, then the fundamental voltage across a tuned circuit in the output has less than a 4.5% peak-to-peak variation:

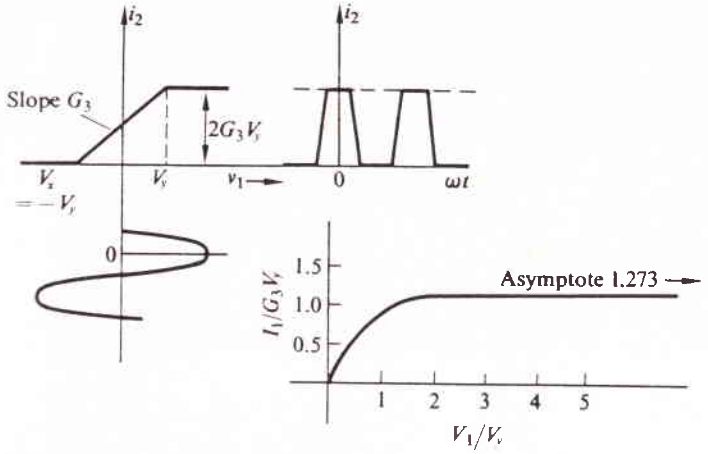


Fig. 4.3–2 Plot of I_1/G_3V_y vs. V_1/V_y.

$$\frac{1.273 - 1.218}{1.24} \times 100 \approx 4.4.$$

Hence on the normal peak basis the output modulation is less than 2.5% or equivalently, we have reduced the AM by more than 30 dB. Such circuitry will be discussed in more detail in Chapter 12.

4.4 SQUARE-LAW CHARACTERISTICS

A square-law voltage in–voltage out characteristic can be approximated by a network of diodes, resistors, and batteries. A square-law voltage in–current out characteristic is approximated quite closely by many field effect transistors (FET) of both the junction and the MOS types operating in their "constant current" region. Figure 4.4–1 shows a typical square-law network for which

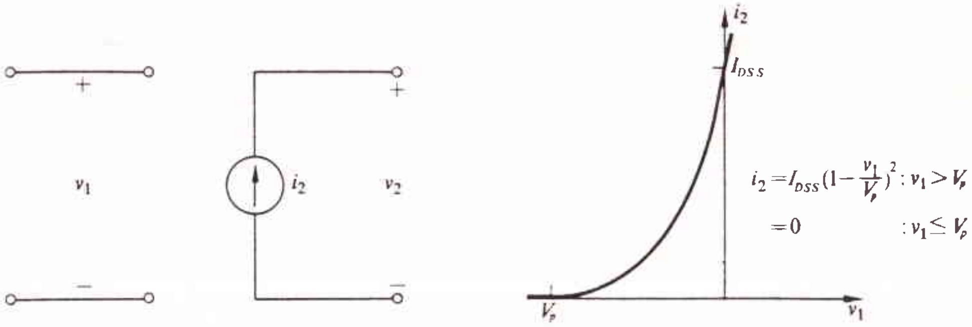

+
v_1
−
i_2
+
v_2
−
i_2
I_{DSS}
$i_2 = I_{DSS}(1-\frac{v_1}{V_p})^2 : v_1 > V_p$
$= 0 \qquad : v_1 \leq V_p$
V_p
v_1

Fig. 4.4–1 Square-law source.

$$i_2 = \begin{cases} I_{DSS}\left(1 - \dfrac{v_1}{V_p}\right)^2 & v_1 > V_p, \\ 0, & v_1 \leq V_p, \end{cases} \tag{4.4–1}$$

where I_{DSS} and V_p are constants similar in notation to those employed with junction-type field effect transistors. Clearly, as v_1 approaches V_p (which is sometimes referred to as the pinch-off voltage), i_2 is reduced to zero. For $v_1 = 0$, $i_2 = I_{DSS}$. In this section V_p is assumed to be negative. Positive values of V_p merely shift the i_2-v_1 characteristic to the right.

Before proceeding further, we should note that when we refer to a square-law characteristic it is not a "true" square-law characteristic, but rather a "half" square-law characteristic. A true square law would be a parabola and therefore a double-valued function of i_2. Since such characteristics do not often exist in nature, we restrict our attention to those formed by one branch of the parabola, as given by Eq. 4.4–1. However, one must be sure that operation is entirely in the square-law region before substituting into the expression

$$i_2 = I_{DSS}\left(1 - \frac{v_1}{V_p}\right)^2.$$

Again, with the square-law characteristic, we find that a square-wave input leads to a square-wave output; hence Eq. (4.2–2) is directly applicable to obtaining the Fourier series of i_2 once the peak value of i_2 is determined.

If, on the other hand, $v_1 = V_b + V_1 \cos \omega t$ and operation is within the square-law region, then

$$i_2 = \frac{I_{DSS}}{V_p^2}(V_x^2 - 2V_xV_1 \cos \omega t + V_1^2 \cos^2 \omega t). \tag{4.4–2}$$

where $V_x = V_p - V_b$. In this case a Fourier series expansion for i_2 has only three terms; that is,

$$i_2(t) = I_0 + I_1 \cos \omega t + I_2 \cos 2\omega t,$$

where

$$I_0 = \frac{I_{DSS}}{V_p^2}\left(V_x^2 + \frac{V_1^2}{2}\right), \tag{4.4–3}$$

$$I_1 = -2\frac{I_{DSS}}{V_p^2} V_x V_1, \tag{4.4–4}$$

$$I_2 = \frac{I_{DSS}}{V_p^2}\frac{V_1^2}{2}. \tag{4.4–5}$$

Thus we find that if V_b is supplied by an ideal voltage source so that V_b is not a function of I_0, then I_1 is a linear function of V_1 and we can define a large-signal average transductance G_m,

$$G_m = \frac{I_1}{V_1} = -2\frac{I_{DSS}}{V_p^2} V_x, \tag{4.4–6}$$

that is independent of the drive voltage. We can define such a large-signal transconductance for any type of characteristic; however, in general, it is not a constant but a function of V_1. When G_m is independent of V_1 and i_2 drives a high-Q_T parallel resonant circuit tuned to the fundamental frequency, then variations in the amplitude of the output voltage are linearly related to variations in V_1; hence, here again, the overall transfer characteristic is linear and the device performs as a *linear* amplifier for AM signals.

The square-law characteristic has the interesting property that the small-signal transconductance g_m at any particular Q-point is equal to the large-signal average transconductance G_m at that same Q-point. To demonstrate this property, we evaluate g_m in the form

$$g_m = \left.\frac{\partial i_2}{\partial v_1}\right|_{v_1 = V_b} = \left.\frac{-2I_{DSS}}{V_p^2}(V_p - v_1)\right|_{v_1 = V_b} = \frac{-2I_{DSS}}{V_p^2} V_x, \tag{4.4–7}$$

which is identical to G_m given by Eq. (4.4–6). A plot of $g_m = G_m$ vs. bias voltage V_b is shown in Fig. 4.4–2. It should be made clear that g_m is an incremental slope at a point on the current-vs.-voltage curve, while $G_m = I_1/V_1$ is a ratio of fundamental output current to fundamental input voltage at a particular operating point.

As an example, suppose that $V_b = V_p/2$, or equivalently $V_x = V_p/2$; then

$$g_m = -\frac{I_{DSS}}{V_p}, \qquad I_1 = -\frac{I_{DSS}}{V_p} V_1, \qquad \text{and} \qquad I_2 = \frac{I_{DSS}}{2}\frac{V_1^2}{V_p}.$$

When V_1 is very small with respect to $|V_p|$, then I_2 vanishes and low-distortion wideband amplification occurs. As V_1 increases, I_2 increases and distortionless wideband amplification is impossible; however, if a narrowband filter is employed, then linear amplification is possible from $V_1 = 0$ to $V_1 = |V_p|/2$, at which point operation exceeds the square-law region of the characteristic and G_m becomes a nonlinear function of V_1.

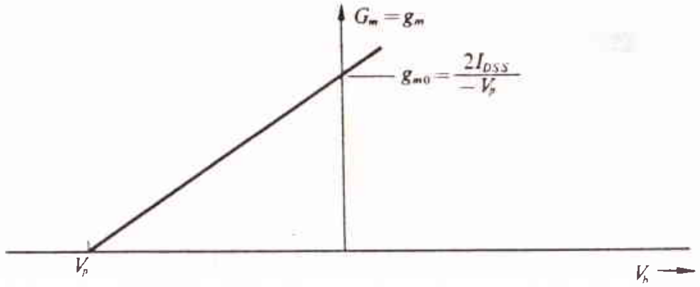


Fig. 4.4–2 Linear transconductances for square-law current characteristic (FET).

In the general case, the second-harmonic distortion in the output current as a function of drive voltage is given by

$$\frac{I_2}{I_1} = -\frac{V_1}{4V_x} = \frac{V_1}{4V_p}\frac{g_{m0}}{g_m}, \tag{4.4–8}$$

where g_{m0} is the value of g_m obtained for $V_b = 0$. Thus, if $V_b = V_p/2$, then 1% distortion arises when $V_1 = |V_p|/50 = |V_b|/25$. This means that linear *small-signal* wideband amplification is restricted to small drive voltages. For example, if $V_p = -4$ V and $V_b = -2$ V, then to keep the distortion below 1% V_1 must be less than 80 mV. As Eq. (4.4–8) indicates, increasing the gain by changing V_b so as to increase g_m

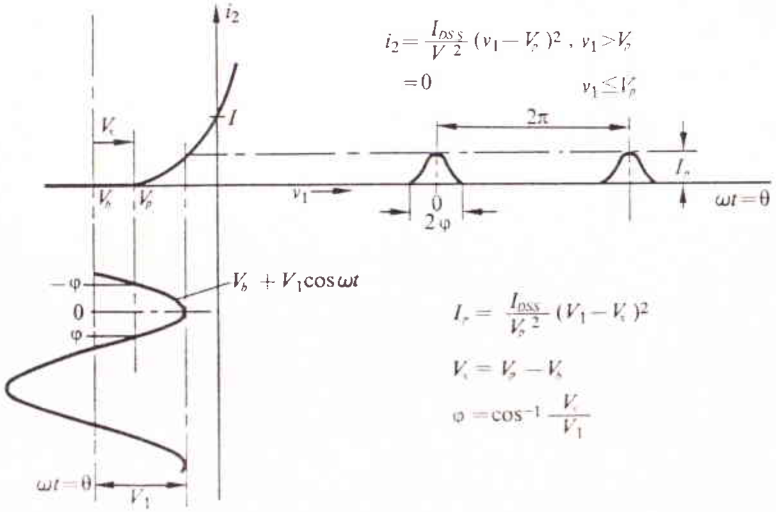

i_2
$i_2 = \frac{I_{DSS}}{V_p^2}(v_1 - V_p)^2, \quad v_1 > V_p$
$= 0 \quad v_1 \leq V_p$
V_x
I
2π
I_p
V_b
V_p
v_1
0
2φ
$\omega t = \theta$
$-\varphi$
0
φ
$V_b + V_1 \cos \omega t$
$I_p = \frac{I_{DSS}}{V_p^2}(V_1 - V_x)^2$
$V_x = V_p - V_b$
$\varphi = \cos^{-1}\frac{V_x}{V_1}$
$\omega t = \theta$
V_1

Fig. 4.4–3 Square-law device driven beyond cutoff.

does decrease the distortion, but not enough to make this type of characteristic well suited to producing wideband amplifiers with large-signal handling capabilities.

In practical devices, the characteristic is likely to depart somewhat from square law at its extreme values; hence, if the above results are to apply, drive voltages must be small enough to stay out of these regions. Even if Eq. (4.4–1) does accurately describe the device, if $v_1 = V_b + V_1 \cos \omega t$ exceeds V_p for some portion of a cycle the above results are not valid. In particular, the device is off for part of the cycle and square law for the remainder of the cycle, as illustrated in Fig. 4.4–3. Thus we see that the current is a periodic train of square-law sine-wave tips of peak amplitude

$$I_p = \frac{I_{DSS}}{V_p^2}(V_1 - V_x)^2.$$

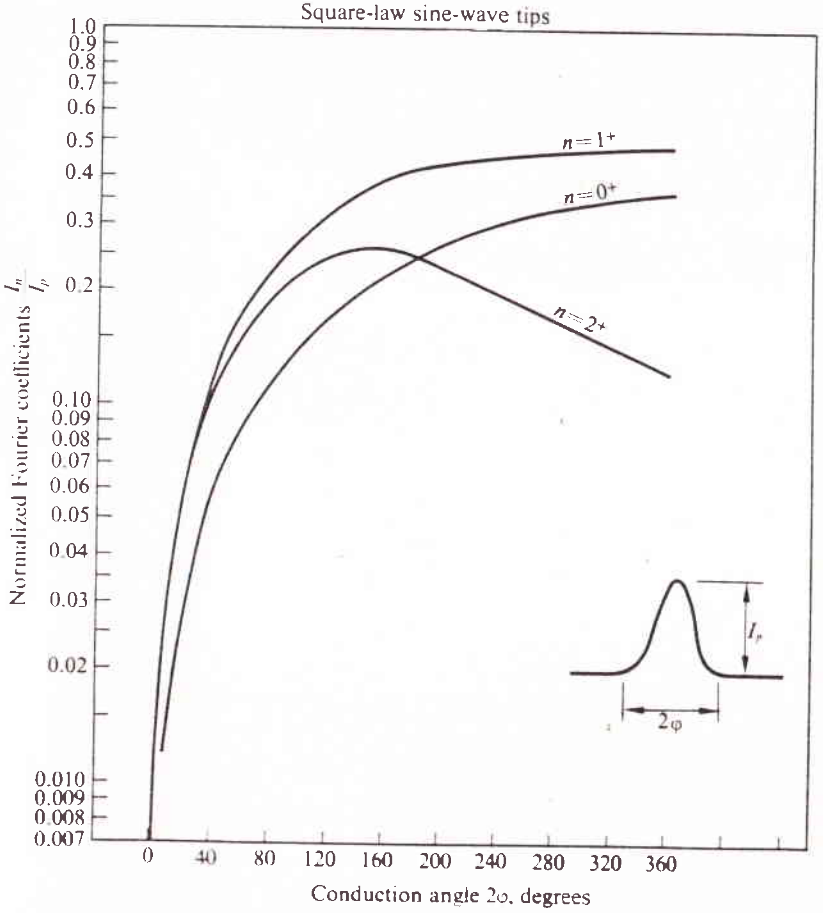


Fig. 4.4–4 Normalized Fourier coefficients of a square-law sine-wave tip pulse train vs. the conduction angle.

In addition, the conduction angle 2ϕ (i.e., the angular portion of the cycle during which i_2 is different from zero) is given by

$$2\phi = 2 \cos^{-1} \frac{V_x}{V_1}.$$

Under this mode of operation the Fourier series for i_2 no longer consists of three terms but rather an infinite number of terms. The coefficients of these terms of the Fourier series as functions of I_p and ϕ are presented in the appendix to this chapter. Figure 4.4–4 presents the normalized values for the first three coefficients in terms of conduction angle 2ϕ.

In addition, Fig. 4.4–5 presents the normalized Fourier coefficients I_n/I_p plotted vs. $V_x/V_1 = \cos \phi$ for $n = 0, 1,$ and 2. Negative values of V_x/V_1 are obtained when

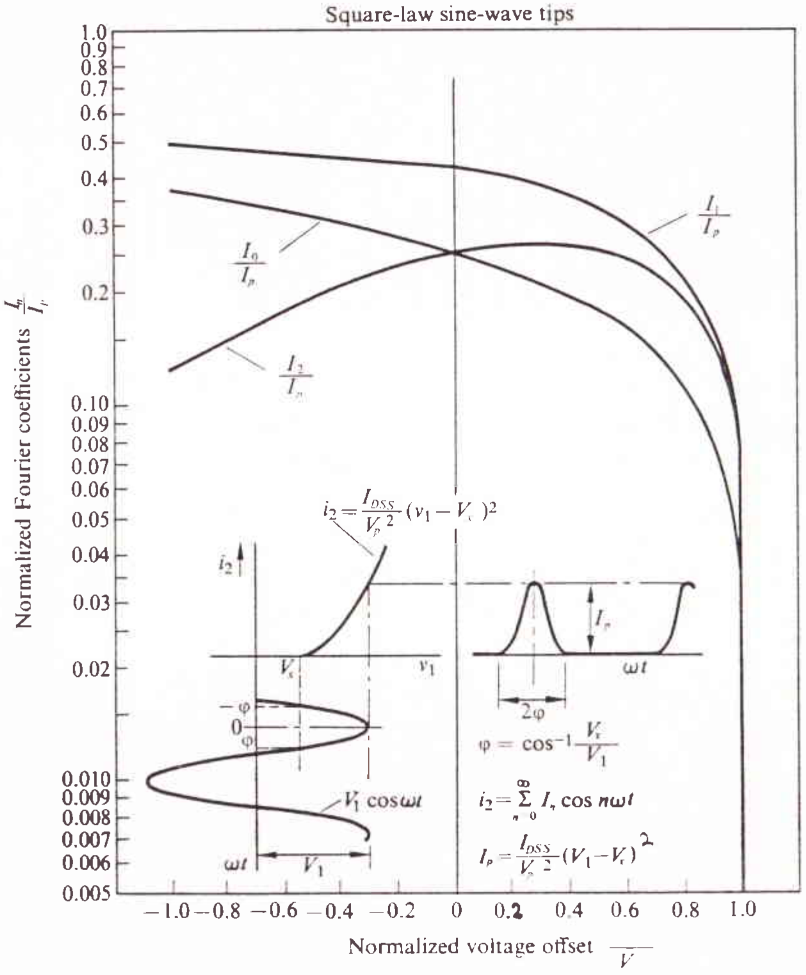


Fig. 4.4–5 Plot I_n/I_p vs. V_x/V_1 for $n = 0, 1, 2$.

the bias voltage V_b lies within the square-law region, or equivalently, when the conduction angle is greater than 180°. The point $V_x/V_1 = -1$ corresponds to complete operation within the square-law region and a conduction angle of 360°; consequently, for values of $V_x/V_1 < -1$, the operation is well within the square-law region and the Fourier coefficients (only three) must be found from Eqs. (4.4–3), (4.4–4), and (4.4–5), which may be manipulated into the equivalent forms

$$I_0 = I_p \frac{\frac{1}{2} + (V_x/V_1)^2}{(1 - V_x/V_1)^2}, \tag{4.4–9}$$

$$I_1 = I_p \frac{-2(V_x/V_1)}{(1 - V_x/V_1)^2}, \tag{4.4–10}$$

$$I_2 = I_p \frac{\frac{1}{2}}{(1 - V_x/V_1)^2}, \tag{4.4–11}$$

which are valid for $V_x/V_1 \leq -1$ and for which, of course,

$$I_p = \frac{I_{DSS}}{V_p^2}(V_1 - V_x).$$

It is of interest to note that for $V_x/V_1 = -1$ substituted into Eqs. (4.4–9), (4.4–10), and (4.4–11), $I_0/I_p = \frac{3}{8}$, $I_1/I_p = \frac{1}{2}$, and $I_2/I_p = \frac{1}{8}$. These are exactly the values shown in Fig. 4.4–5 for $V_x/V_1 = -1$; hence the Fourier coefficients I_n vary continuously as V_x/V_1 decreases beyond -1.

These square-law sine-wave tip characteristics will be quite useful in designing self-limiting field effect transistor oscillators, as well as in explaining the action in large-signal FET RF amplifiers.

4.5 THE EXPONENTIAL CHARACTERISTIC

A very good approximation to the current emitted across a forward-biased P-N junction in a junction diode or transistor is

$$i_2 = I_S e^{v_1 q/kT}, \tag{4.5–1}$$

where kT/q is approximately 26 mV for $T = 300°\text{K}$. If we assume that the transistor alpha is independent of current and ignore for the present all internal resistive voltage drops, then by choosing I_S appropriately we can use Eq. (4.5–1) to represent emitter, collector, or base currents. (The equation is not "exact" in the reverse direction; however, for our purposes this "nonexactness" is a third-order effect.)

A typical voltage in–current out exponential characteristic is obtained from the circuit of Fig. 4.5–1. If $v_1 = V_b + v(t)$ is applied to the device, $i_2(t)$ has the form $I_{dc} + i(t)$, where I_{dc} and $i(t)$ are functions of both V_b and $v(t)$. In our previous examples we assumed that V_b and $v(t)$ were supplied from independent sources and derived our results in terms of these quantities. In this case, however, we assume that I_{dc} and $i(t)$ are supplied from independent sources and present the results in terms of these quantities. The reason for this choice is apparent when we realize that in a well-designed transistor circuit we control the bias (or average) current and the ac junction voltage; hence these are the quantities we know.

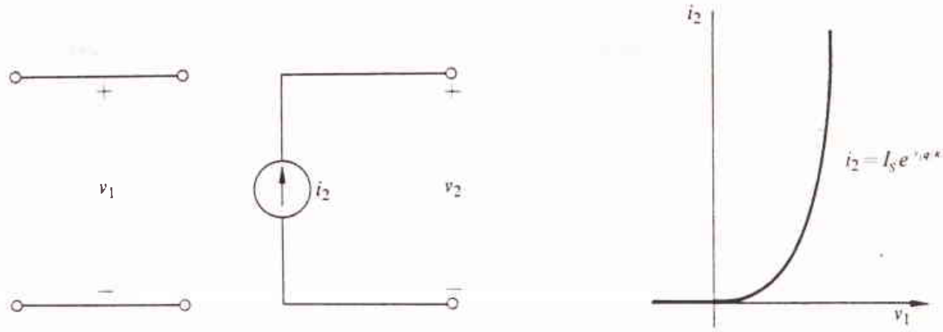


Fig. 4.5–1 Exponentially controlled current source.

With $v_1(t) = V_b + v(t)$ applied to the circuit of Fig. 4.5–1, $i_2(t)$ is given by

$$i_2 = I_S e^{qV_b/kT} e^{qv(t)/kT}.$$

For small-signal operation, i.e., values of $v(t)$ for which $qv(t)/kT \ll 1$, $i_2(t)$ reduces to the form

$$\begin{aligned} i_2 &= I_S e^{qV_b/kT}\left[1 + \frac{qv(t)}{kT}\right] \\ &= I_{dc} + i(t), \end{aligned} \tag{4.5–2}$$

where $I_{dc} = I_S e^{qV_b/kT}$, and $i(t) = qI_{dc}v(t)/kT$. Clearly $i(t)$ and $v(t)$ are linearly related by a small-signal transconductance

$$g_m = \frac{v(t)}{i(t)} = \frac{qI_{dc}}{kT}$$

which is a linear function of bias current. We note that small-signal linear "gain" control can be achieved by controlling the bias current I_{dc}. This linear gain control will be seen to have many useful applications.

Square-Wave Input

If $v(t)$ is a symmetrical square wave of peak amplitude V_1, it is apparent that i_2 is also a square wave with an upper level

$$I_M = I_S e^{qV_b/kT} e^{qV_1/kT}$$

and a lower level

$$I_m = I_S e^{qV_b/kT} e^{-qV_1/kT}.$$

Thus the average value of i_2 is given by

$$I_{dc} = \frac{I_M + I_m}{2} = I_S e^{qV_b/kT} \cosh x \tag{4.5–3}$$

and the peak-to-peak value of i_2 is given by

$$I_{pp} = I_M - I_m = 2I_S e^{qV_b/kT} \sinh x, \tag{4.5–4}$$

where $x = qV_1/kT$ normalizes the input square-wave amplitude to 26 mV (at $T = 300°$ K). By combining Eqs. (4.5–3) and (4.5–4), we obtain

$$I_{pp} = 2I_{dc} \tanh x. \tag{4.5–5}$$

The quantity $I_{pp}/2I_{dc}$ is plotted in Fig. 4.5–2 vs. x.

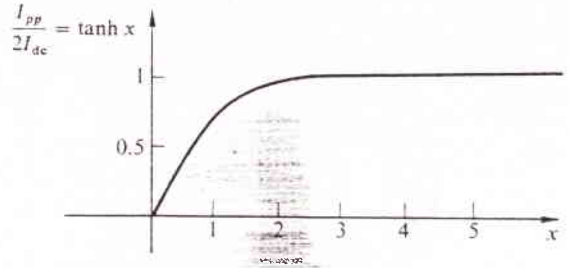


Fig. 4.5–2 Plot of $I_{pp}/2I_{dc}$ vs. x.

For $x \ll 1$, $\tanh x \approx x$ and Eq. (4.5–5) reduces to

$$\frac{I_{pp}}{2} = I_{dc}x = \frac{qI_{dc}}{kT} V_1 = g_m V_1, \tag{4.5–6}$$

which, of course, is just the small-signal relationship between the peak amplitude of the output and input square waves. For large values of x ($x > 2.5$), $\tanh x$ is within 1 % of unity and

$$\frac{I_{pp}}{2} = I_{dc}. \tag{4.5–7}$$

This relationship makes sense. For large values of V_1 the negative portion of the input square wave effectively cuts off i_2; hence i_2 varies from zero to I_{pp} and thus has an average value of $I_{pp}/2$.

It is apparent that, for large values of x, the output square-wave amplitude is independent of the input square-wave amplitude and therefore the circuit functions as a limiter. Conversely, for $x \ll 1$ the circuit performs as a linear amplifier.

Regardless of the value of x, however, the peak amplitude of the output square wave is linearly related to I_{dc}; consequently, by controlling I_{dc} we can linearly amplitude-modulate the output square wave. If, in addition, we pass $i_2(t)$ through a high-Q tuned circuit, we obtain an output voltage which is an amplitude-modulated sine wave. In particular, if $i_2(t)$ drives a high-Q parallel RLC circuit which has resistance R_T and is tuned to the fundamental of i_2, then the output voltage $v_2(t)$ is given by (cf. Eq. 4.2–2)

$$v_2(t) = \frac{4I_{dc}R_T}{\pi} \tanh x \cos \omega t. \tag{4.5–8}$$

It is assumed here that if I_{dc} is a function of time, the bandwidth of the tuned circuit is sufficient to pass the AM sideband information.

Sine-Wave Input

If $v_1(t) = V_b + V_1 \cos \omega t$, then

$$\begin{aligned} i_2(t) &= I_S e^{qV_b/kT} e^{qV_1 \cos \omega t/kT} \\ &= I_S e^{qV_b/kT} e^{x \cos \omega t}, \end{aligned}$$

where again $x = qV_1/kT$. For $\omega t = 0, 2\pi, 4\pi, \ldots, i_2(t)$ attains its peak value of $I_p = I_S e^{qV_b/kT} e^x$. To observe the form of $i_2(t)$ for other values of t we normalize $i_2(t)$ to I_p and plot

$$\frac{i_2(t)}{I_p} = \frac{e^{x \cos \omega t}}{e^x} = W_x(t)$$

vs. ωt with x as a parameter, as shown in Fig. 4.5–3. Note that for small values of x the output current is almost cosinusoidal, as would be expected; however, as x increases, the output current becomes pulselike in form.

Because of the exponential nature of i_2, it is not possible to define a conduction angle for these pulses in a conventional sense: however, we may define a fictitious conduction angle as the angular portion of the cycle for which $i_2(t)/I_p \geq 0.05$. This

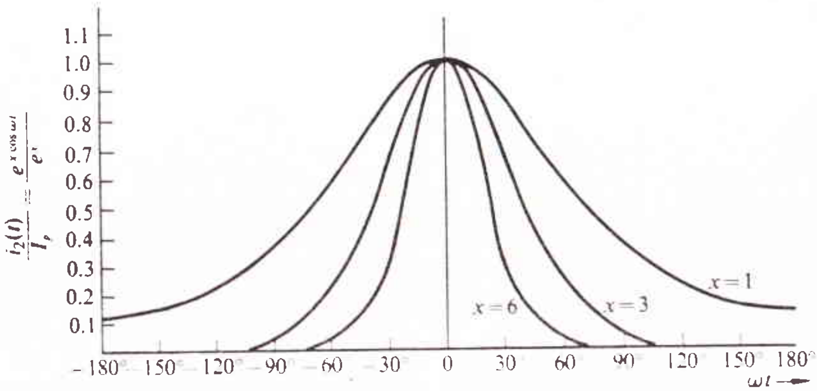


Fig. 4.5–3 Plot of $i_2(t)/I_p$ vs. ωt with $x = qV_1/kT = V_1/(26 \text{ mV})$ as a parameter.

conduction angle may be defined as 2ϕ, where ϕ is the solution of the equation

$$\frac{e^{x \cos\phi}}{e^x} = 0.05, \tag{4.5–9}$$

or equivalently,

$$\phi = \cos^{-1}\left(1 + \frac{\ln 0.05}{x}\right) \approx \cos^{-1}\left(1 - \frac{3}{x}\right). \tag{4.5–10}$$

A plot of this conduction angle vs. x is given in Fig. 4.5–4.

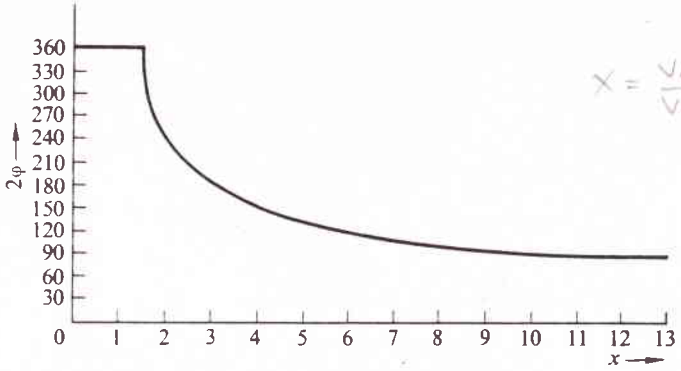


Fig. 4.5–4 Plot of conduction angle 2ϕ vs. x.

$X = \frac{V_1}{V_T}$

The pulse shape is a function only of the normalized ac input voltage (x), and not of the bias level. This, of course, is a basic property of an exponential characteristic.

The current $i_2(t) = I_p e^{x \cos\omega t}/e^x$ may be expanded in a Fourier series of the form

$$i_2(t) = \sum_{n=0}^{\infty} C_n \cos n\omega t,$$

where

$$C_0 = \frac{I_p}{e^x}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{x\cos\theta}\, d\theta\right) = \frac{I_p}{e^x} I_0(x), \tag{4.5–11}$$

$$C_n = \frac{2I_p}{e^x}\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{x\cos\theta} \cos n\theta,\, d\theta\right) = \frac{2I_p}{e^x} I_n(x), \tag{4.5–12}$$

and $I_n(x)$ is a modified Bessel function of order n and argument x. The modified Bessel functions should not frighten the reader, since they represent merely the tabulated results of the numerical integrations of the integrals of Eqs. (4.5–11) and (4.5–12) with x as a parameter. Various useful properties of the modified Bessel functions, including their tabulation, are presented in the Appendix at the back of the book.

By noting that

$$C_0 = \frac{I_p}{e^x} I_0(x) = I_{dc}, \tag{4.5–11a}$$

we can write $i_2(t)$ as a function of I_{dc} in the form

$$i_2(t) = I_{dc}\left[1 + \sum_{n=1}^{\infty} \frac{2I_n(x)}{I_0(x)} \cos n\omega t\right]. \tag{4.5–13}$$

Table 4.5–1 Tabulation of modified Bessel function ratios vs. x

x	$\frac{2I_1(x)}{I_0(x)}$	$\frac{2I_2(x)}{I_0(x)}$	$\frac{2I_3(x)}{I_0(x)}$	$\frac{I_2(x)}{I_1(x)}$
0.0	0.0	0.0	0.0	0.0
0.1	0.0999	0.0024	—	0.024
0.5	0.4850	0.0600	0.0050	0.124
1.0	0.8928	0.2144	0.0350	0.240
2.0	1.3955	0.6045	0.1866	0.433
3.0	1.6200	0.9200	0.3933	0.568
5.0	1.7868	1.2853	0.7585	0.719
7.0	1.8511	1.4711	1.0104	0.795
10.0	1.8972	1.6206	1.2490	0.854
14.0	1.9272	1.7247	1.4344	0.895
20.0	1.9493	1.8051	1.5883	0.926

Table 4.5–1 presents data for $2I_1(x)/I_0(x)$, $2I_2(x)/I_0(x)$, $2I_3(x)/I_0(x)$, and $I_2(x)/I_1(x)$ all vs. x. Figure 4.5–5 presents a plot of these same data. [These data have been calculated from the values of $I_n(x)$ given in the Appendix at the back of the book.] As shown in the Appendix, for small values of x, $I_0(x) \to 1$, $I_1(x) \to x/2$, and $I_n(x) \to 0$

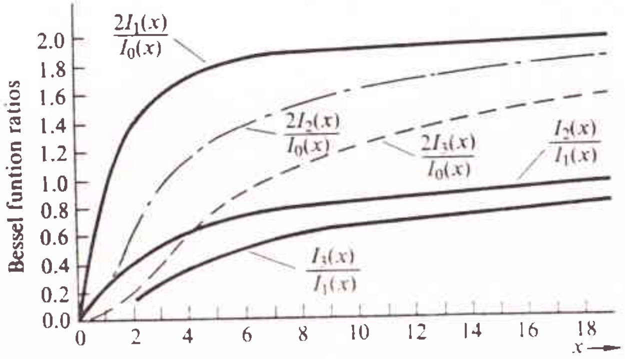


Fig. 4.5–5 Modified Bessel function ratios vs. $x = qV_1/kT$

(actually some power of x greater than unity); hence, when x approaches zero,

$$i_2(t) \approx I_{dc}(1 + x \cos \omega t) = I_{dc} + g_m V_1 \cos \omega t, \tag{4.5–14}$$

which is the expected small-signal response.

For large values of x, the pulse width of i_2 becomes narrow and the fundamental component is seen to approach twice the dc value, as expected; thus again the exponential characteristic is useful as a limiter so long as the minimum amplitude does not cause x to drop below say 5. For example, with a normalized carrier amplitude of $x = 10$, 50% amplitude modulation in the input ($m = 0.5$) causes less than 4% output amplitude modulation ($m_o = 0.04$).

When x is large, $I_0(x)$ approaches $e^x/\sqrt{2\pi x}$ (this equality is within 6% for $x \geq 3$ and within 3% for $x \geq 5$). Hence from Eq. (4.5–11a) the relation between the dc component and the peak current for $x \geq 3$ is given by

$$I_p \approx I_{dc}\sqrt{2\pi x}.$$

Thus in a circuit in which I_{dc} is held constant at 1 mA (and in which resistive drops are negligible), a peak sinusoidal base-emitter voltage drive of 260 mV will cause a current peak of 7.95 mA, while the fundamental component will have a peak value of only 1.90 mA.

Figure 4.5–5 conveys all sorts of additional interesting information about the exponential characteristic. If I_{dc} is presumed constant, so that it is not a function of V_1, then the $2[I_1(x)/I_0(x)]$ curve indicates that a narrowband "linear" amplifier for the I_1 component is feasible only for values of x less than unity. From the $I_2(x)/I_1(x)$ curve one sees that, for wideband linear amplifiers, x must be less than 0.1 just to keep the second-harmonic distortion below 2.5%. [Note as $x \to 0$, $I_2(x)/I_1(x) \to x/4$: hence $x = 0.1$ implies $I_2(x)/I_1(x) = 0.025$.] Therefore, such a characteristic is useful only in an absolutely linear fashion when input voltages are below 2.5 mV. (The schemes that one employs to design amplifiers capable of handling large-signal inputs without distortion will be discussed in subsequent sections.)

When the exponential characteristic is followed by a tuned circuit which extracts the fundamental component of $i_2(t)$, it is again convenient to define a large-signal average fundamental transconductance $G_m(x)$ as

$$G_m(x) = \frac{I_1}{V_1} = \frac{I_{dc}}{V_1}\frac{2I_1(x)}{I_0(x)} = \frac{qI_{dc}}{kT}\frac{2I_1(x)}{xI_0(x)}. \tag{4.5–15}$$

With this definition the output voltage $v_2(t)$ can be written as

$$v_2(t) = -G_m(x)V_1 R_T \cos \omega t, \tag{4.5–16}$$

where R_T is the resistance of the tuned circuit at resonance. If I_{dc} is independent of V_1, then

$$I_{dc} = I_{dc}\big|_{V_1=0} \qquad \text{and} \qquad \frac{qI_{dc}}{kT} = g_m.$$

Thus $G_m(x)$ as given by Eq. (4.5–15) reduces to

$$G_m(x) = g_m \frac{2I_1(x)}{xI_0(x)}. \tag{4.5–17}$$

Table 4.5–2 presents values for $G_m(x)/g_m = 2I_1(x)/xI_0(x)$ vs. x. In addition, these data are plotted in Fig. 4.5–6.

Table 4.5–2 Tabulation of $2I_1(x)/xI_0(x)$ vs. x

x	$\dfrac{2I_1(x)}{xI_0(x)} = \dfrac{G_m(x)}{g_m}$
0.0	1.0
0.2	0.995
0.5	0.970
1.0	0.893
2.0	0.698
3.0	0.540
4.0	0.432
5.0	0.357
6.0	0.304
7.0	0.264
8.0	0.234
9.0	0.210
10.0	0.190
15.0	0.129
20.0	0.0975

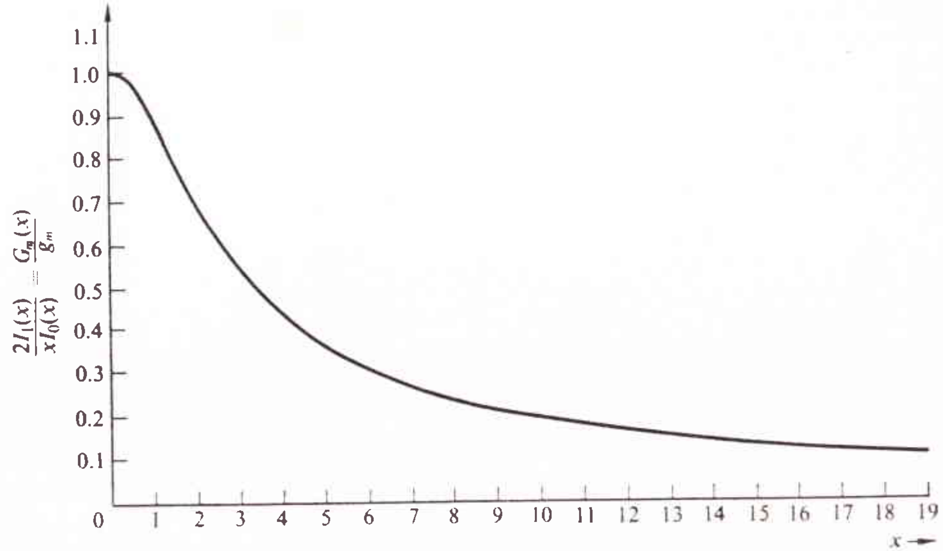


Fig. 4.5–6 Plot of $2I_1(x)/xI_0(x)$ vs. x.

By the time $x = 1$, $G_m(x)/g_m$ is already down by 1 dB; hence we again observe that "linear" narrowband AM amplifiers require $|v_1|_{\max} < 26$ mV. However, the drop in $G_m(x)$ with increasing drive is useful in some circuits. For example, it allows the output amplitude of a sine-wave oscillator to stabilize. As the oscillations begin to grow, the loop gain decreases to the point where a further increase in the oscillation amplitude is not possible.

We should note finally that $i_2(t)$, given by Eq. (4.5–13), is directly proportional to I_{dc}; hence, here again, by controlling I_{dc} we can achieve multiplication or amplitude modulation independent of the value of V_1 (or x). In Chapter 8 we shall consider in detail how this control is accomplished.

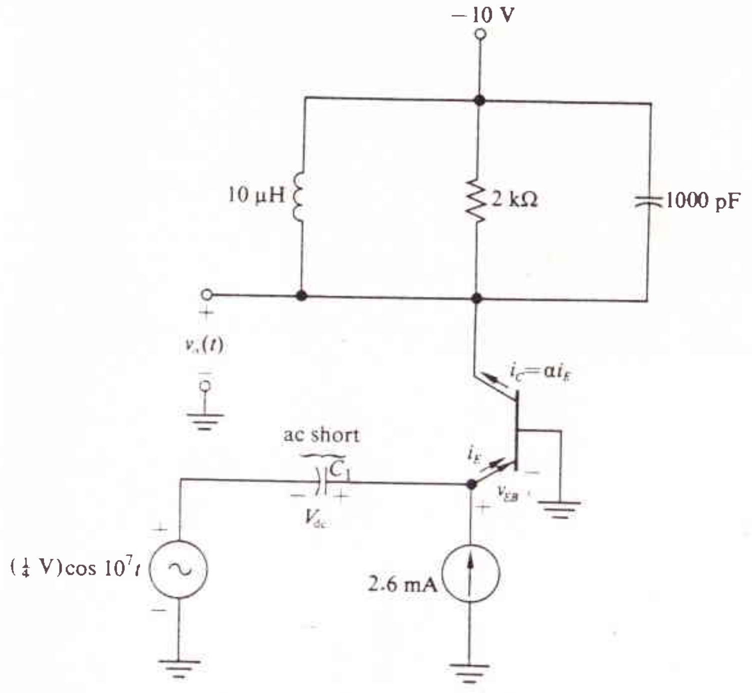


Figure 4.5–7

Example 4.5–1 For the narrowband amplifier shown in Fig. 4.5–7, determine an expression for $v_o(t)$. Also determine the peak emitter current and its conduction angle.

Solution. The base-emitter voltage consists of the input cosine plus whatever dc voltage V_{dc} builds up across the capacitor C_1; hence

$$v_{EB} = V_{dc} + \tfrac{1}{4}\cos 10^7 t$$

and

$$\begin{aligned} i_E &= I_{ES}e^{qV_{dc}/kT}e^{9.6\cos 10^7 t} \\ &= I_p\frac{e^{9.6\cos 10^7 t}}{e^{9.6}} \\ &= \frac{I_p}{e^{9.6}}I_0(9.6)\left[1 + \frac{2I_1(9.6)}{I_0(9.6)}\cos 10^7 t + \cdots +\right]. \end{aligned}$$

The 2.6-mA current source must equal the average value of i_E; therefore,

$$\frac{I_p I_0(9.6)}{e^{9.6}} = I_{dc} = 2.6 \text{ mA}$$

and the fundamental component I_{E1} is equal to

$$\begin{aligned} I_{E1} &= (2.6 \text{ mA})\frac{2I_1(9.6)}{I_0(9.6)} \\ &= (2.6)(1.88) \text{ mA} = 4.88 \text{ mA}. \end{aligned}$$

Since the output-tuned circuit resonates at $\omega_0 = 1/\sqrt{LC} = 10^7$ rad/sec and has $Q_T = \omega_0 RC = 20$, only the fundamental component of i_E contributes to the output; therefore, if $\alpha \approx 1$,

$$v_o = (-10 \text{ V}) + (2 \text{ k}\Omega)(\alpha)(4.88 \text{ mA})\cos 10^7 t = (-10 \text{ V}) + (9.8 \text{ V})\cos 10^7 t.$$

In addition, since $I_p = I_{dc}e^{9.6}/I_0(9.6)$, then by interpolating from Table A-2, or using the relationship $I_p \approx \sqrt{2\pi x}I_{dc}$, we obtain

$$I_p = \frac{2.6 \text{ mA}}{0.13} = 20 \text{ mA}.$$

Also, from Eq. (4.5–10) and by noting that $x = 9.6$, we obtain a conduction angle of 96°.

As an alternative approach to obtaining v_o, we observe that the small-signal transconductance g_m is

$$g_m = \alpha g_{in} = \frac{\alpha q I_{dc}}{kT} \approx 0.1 \mho.$$

Consequently, the large-signal average transconductance (cf. Fig. 4.5–6) is given by

$$G_m = (0.197)(0.1)\mho = 0.0197\mho.$$

Therefore, the output voltage has the value

$$v_o(t) = V_{CC} + G_m(2 \text{ k}\Omega)[(\tfrac{1}{4} \text{ V})\cos 10^7 t] = (-10 \text{ V}) + (9.8 \text{ V})\cos 10^7 t,$$

which, of course, agrees with the previous evaluation.

6 THE DIFFERENTIAL CHARACTERISTIC

two transistors with characteristics of the type outlined by Eq. (4.5–1) are con-
ected in a differential configuration so that their total emitter current, I_k, is supplied
y a constant current source (normally another transistor), then a very useful and
asily integrable circuit results. The differentially connected circuit is shown in
ig. 4.6–1. This circuit has a distinct nonlinear characteristic that should be added
our collection of useful nonlinearities.

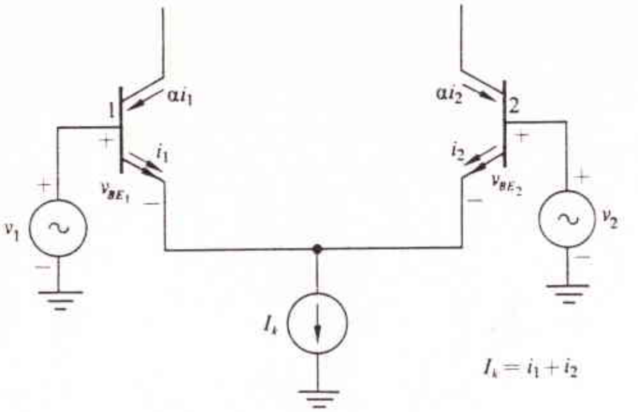


Fig. 4.6–1 Differential configuration.

To obtain an expression for the nonlinear characteristic of the differentially
onnected circuit, we note that

$$i_1 = I_{S1}e^{v_{BE1}q/kT}, \qquad i_2 = I_{S2}e^{v_{BE2}q/kT}, \tag{4.6–1}$$

nd

$$v_{BE1} - v_{BE2} = v_1 - v_2.$$

herefore, if $I_{S1} = I_{S2}$ (which should be true if both transistors are integrated on
he same "chip" and are identical in size and construction), then

$$\frac{i_1}{i_2} = e^{(v_1 - v_2)q/kT}. \tag{4.6–2}$$

we now set $i_1 + i_2 = I_k$, we obtain

$$i_2 = \frac{I_k}{1 + e^z} \qquad \text{and} \qquad i_1 = \frac{I_k}{1 + e^{-z}}, \tag{4.6–3}$$

here $z = (v_1 - v_2)q/kT$. The normalized nonlinear characteristics i_1/I_k and
$_2/I_k$ vs. z are tabulated in Table 4.6–1 and plotted in Fig. 4.6–2.

Table 4.6–1

z	i_1/I_k	i_2/I_k	$v_1 - v_2$
+2.94	0.95	0.05	+77 mV
+2.20	0.90	0.10	+57 mV
+1.10	0.75	0.25	+28.6 mV
+0.405	0.60	0.40	+10.4 mV
0.000	0.50	0.50	0.0
−0.405	0.40	0.60	−10.4 mV
−1.10	0.25	0.75	−28.6 mV
−2.20	0.10	0.90	−57 mV
−2.94	0.05	0.95	−77 mV

It is of interest to note that both i_1 and i_2 possess odd symmetry about their average value $I_k/2$. Specifically,

$$i_2 - \frac{I_k}{2} \equiv -i(z) = \frac{I_k}{2}\left(\frac{1 - e^z}{1 + e^z}\right) = -\frac{I_k}{2}\tanh\frac{z}{2}, \tag{4.6–4}$$

from which we observe that $i(z) = -i(-z)$. With the aid of Eq. (4.6–4) we can write i_2 in the alternative form

$$\frac{i_2}{I_k} = \frac{1}{2}\left(1 - \tanh\frac{z}{2}\right) \tag{4.6–5}$$

and, in a similar fashion, we write

$$\frac{i_1}{I_k} = \frac{1}{2}\left(1 + \tanh\frac{z}{2}\right). \tag{4.6–6}$$

If this circuit is operated about its point of symmetry ($z = 0$), as it normally is, then the dc component in either current remains at $I_k/2$ for all symmetrical driving signals ($v_1 - v_2$). If, in addition, the driving signals are periodic, no even harmonics are generated. In particular, if $v_1 - v_2$ is a square wave of zero average value and peak amplitude V_1, then i_1 is a square wave in phase with the input and i_2 is a square

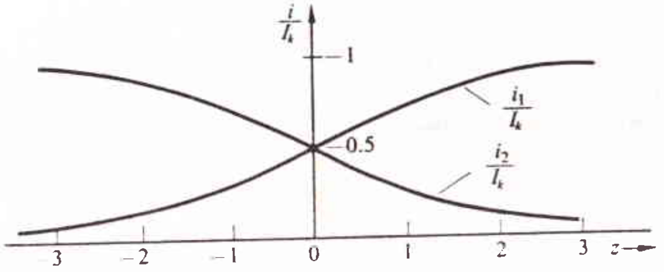


Fig. 4.6–2 Differential characteristics plotted vs. $z = (v_1 - v_2)q/kT$.

wave 180° out of phase with the input. Both i_1 and i_2 are symmetric about $I_k/2$ and have peak values which can be determined directly from Table 4.6–1 or Fig. 4.6–2. For example, if $V_1 = 57\text{ mV}$ ($T = 300°\text{ K}$), then the peak-to-peak value of both i_1 and i_2 is equal to $0.8I_k$.

For small values of z ($z \ll 1$ or equivalently $|v_1 - v_2| \ll 26\text{ mV}$), we may approximate $\tanh(z/2)$ by $z/2$ to obtain

$$i_2 = \frac{I_k}{2}\left(1 - \frac{z}{2}\right) = \frac{I_k}{2} - \frac{g_{\text{in}}}{2}(v_1 - v_2)$$

and $$\tag{4.6–7}$$

$$i_1 = \frac{I_k}{2}\left(1 + \frac{z}{2}\right) = \frac{I_k}{2} + \frac{g_{\text{in}}}{2}(v_1 - v_2),$$

where $g_{\text{in}} = qI_k/2kT$. Since g_{in} is just the small-signal conductance "seen" looking into the emitter of transistor 1 or 2 with the base grounded and with an emitter current of $I_k/2$, it is clear that Eq. (4.6–7) represents simply the expected small-signal transfer for a differential pair of transistors. In addition, the small-signal transconductance

$$g_m = \frac{\alpha i_{2\text{ac}}}{v_1 - v_2} = \frac{-\alpha i_{1\text{ac}}}{v_1 - v_2}$$

is given by

$$g_m = \frac{\alpha g_{\text{in}}}{2}, \tag{4.6–8}$$

where α is the ratio of collector to emitter current.

If now $v_1 - v_2 = V_1 \cos\omega t$ and we again define $x = qV_1/kT$, then

$$i_1 = \frac{I_k}{2} + i \qquad \text{and} \qquad i_2 = \frac{I_k}{2} - i, \tag{4.6–9}$$

where

$$i = \frac{I_k}{2}\tanh\left(\frac{x}{2}\cos\omega t\right).$$

A sketch of i vs. ωt, shown in Fig. 4.6–3, indicates how the waveform varies with x.

It is quite apparent that the Fourier series of i contains no even harmonics; hence i may be expanded in the form

$$i = I_1\cos\omega t + I_3\cos 3\omega t + I_5\cos 5\omega t + \cdots$$

$$= I_k \sum_{n=1}^{\infty} a_{2n-1}(x)\cos(2n-1)\omega t, \tag{4.6–10}$$

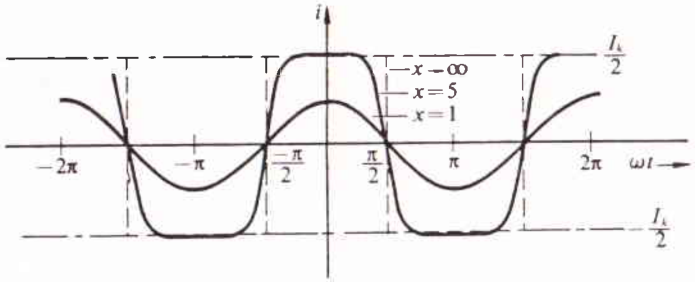


Fig. 4.6–3 Sketch of i vs. ωt for several values of x.

where

$$a_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi}\left[\tfrac{1}{2}\tanh\left(\frac{x}{2}\cos\theta\right)\right]\cos n\theta\, d\theta = \frac{I_n}{I_k}.$$

The coefficients $a_n(x)$ do not appear to be expandable in terms of tabulated functions; however, it is a straightforward numerical matter to determine the Fourier coefficients a_n for the fundamental and the first several harmonics. Table 4.6–2 presents $a_n(x) = I_n/I_k$ for $n = 1, 3, 5$.

Table 4.6–2 Tabulation of I_n/I_k vs. x for $n = 1, 3, 5$

x	$a_1(x) = I_1/I_k$	$a_3(x) = I_3/I_k$	$a_5(x) = I_5/I_k$
0.0	0.0000	0.0000	0.0000
0.5	0.1231	—	—
1.0	0.2356	−0.0046	—
1.5	0.3305	−0.0136	—
2.0	0.4058	−0.0271	—
2.5	0.4631	−0.0435	0.00226
3.0	0.5054	−0.0611	0.0097
4.0	0.5586	—	—
5.0	0.5877	−0.1214	0.0355
7.0	0.6112	−0.1571	0.0575
10.0	0.6257	−0.1827	0.0831
∞	0.6366	−0.2122	+0.1273

Figure 4.6–4 plots $a_1(x)$, $a_3(x)$, and $a_5(x)$ vs. x and compares the fundamental components from the differential characteristic with those of a truncated linear characteristic (Fig. 4.3–1b) that has the same slope at the point of symmetry and the same asymptotic values. The values for $x = \infty$ are obtained from Eq. (4.2–2) by noting that, for very large drive voltages, the currents i_1 and i_2 must have the

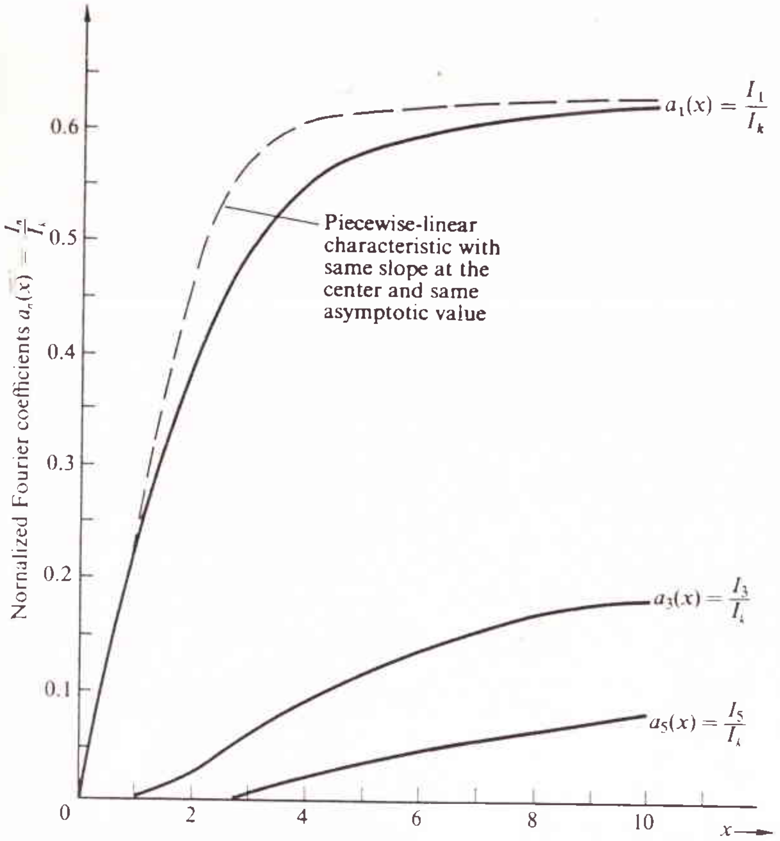


Fig. 4.6–4 Harmonic content from sinusoidally excited differential characteristic.

form of square waves. In addition, for small values of x we may expand $\frac{1}{2}\tanh[(x/2)\cos\omega t]$ in a MacLauren expansion of the form

$$\frac{1}{2}\tanh\left(\frac{x}{2}\cos\omega t\right) = \frac{1}{2}\left(\frac{x}{2}\cos\omega t - \frac{x^3}{24}\cos^3\omega t\right)$$
$$= \frac{1}{2}\left[\left(\frac{x}{2} - \frac{x^3}{32}\right)\cos\omega t - \frac{x^3}{96}\cos 3\omega t\right].$$

Thus, for small x, $a_1(x) \approx \frac{1}{4}(x - x^3/16)$ and $a_3(x) \approx x^3/192$. Clearly, then, $a_1(x)$ is a linear function of x, provided that $x^2/16 \ll 1$. Specifically, if $x \leq 0.63$, the nonlinear term in $a_1(x)$ is less than 2.5% of the linear term; that is, $x^2/16 \leq 0.025$. It is also apparent that for $x \leq 1$, $a_3(x)/a_1(x) \leq 0.02$; consequently, the limits on linearity are determined not by the presence of the third-harmonic current component but rather by the presence of the nonlinear term in $a_1(x)$. Hence broadband as well as narrowband linear amplification is possible for values of $|v_1 - v_2| \leq 16$ mV. This increase in input drive by a factor of 6.3 over the single-transistor broadband

amplifier is due primarily to the symmetry of the differential characteristic and the resultant absence of even (principally second) harmonics from the output. Intuitively what we are doing is compensating for the nonlinear input characteristic of a single transistor by using a second transistor as a nonlinear emitter impedance.

When a tuned circuit which extracts the fundamental component of current is placed in the collector of either transistor, it is convenient to again define a large-signal average fundamental transconductance as

$$G_m(x) = \frac{\alpha I_1}{V_1} = \frac{\alpha q I_k}{kT}\frac{I_1}{x I_k} = 2\alpha g_{in}\frac{a_1(x)}{x} = g_m\frac{4a_1(x)}{x},$$

where g_m is the small-signal transconductance defined by Eq. (4.6–8). With this definition of $G_m(x)$, if a high Q_T tuned circuit with a resistance R_L at resonance is placed in the collector of transistor 2, the tuned-circuit voltage v_T is given by

$$v_T(t) = G_m(x)V_1R_L\cos\omega t. \tag{4.6–11}$$

A negative sign would be required if the tuned circuit were placed in the collector of transistor 1. A plot of $4a_1(x)/x = G_m(x)/g_m$ is given in Fig. 4.6–5.

Finally, it should be noted that the differential characteristic has many of the same properties as the exponential characteristic. For small input signals the differentially connected circuit operates as a linear amplifier, for large input signals,

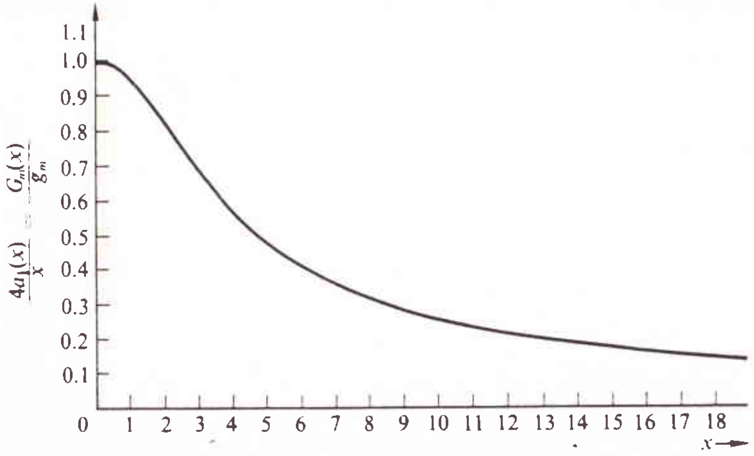


Fig. 4.6–5 Plot of $G_m(x)/g_m = 4a_1(x)/x$ vs. x.

I_1/I_k approaches a constant and the circuit functions as a limiter. In addition, i_1 and i_2 are directly proportional to I_k; therefore, linear gain control or amplitude modulation can be achieved by controlling I_k. In practice, as we shall see in Chapter 5, it turns out to be more effective to control I_k for the differential pair than I_{dc} for the single transistor.

4.7 OTHER GRADUAL NONLINEARITIES—PENTODES

Many practical devices have current-source-like outputs in which the input voltage–output current relationship is neither exponential nor square law. In many cases, this relationship takes the form

$$i_2 = \begin{cases} I_{SS}\left(1 - \dfrac{v_1}{V_{co}}\right)^n & v_1 > V_{co}, \\ 0, & v_1 \le V_{co}. \end{cases} \tag{4.7–1}$$

In this equation I_{SS} is the current flowing when $v_1 = 0$, while V_{co} is a cutoff voltage (that is, $i_2 = 0$ when $v_1 = V_{co}$).

In a pentode, for example, for which the plate-to-cathode voltage v_B is kept above the knee of the v_B-i_B characteristic, and for which the screen-to-cathode voltage v_S and suppressor-to-cathode voltage v_{SP} are held constant, the plate current i_B may be related to the control grid-to-cathode voltage v_G by the relationship

$$i_B = I_{SS}\left(1 - \frac{v_G}{V_{co}}\right)^{3/2}. \tag{4.7–2}$$

Figure 4.7–1 shows a typical set of v_G-i_B characteristics for a pentode with the screen voltage as a parameter, along with the corresponding v_B-i_B characteristic. It should be noted that both I_{SS} and V_{co} vary with the screen grid voltage. In particular, for the pentode shown, with $v_S = 150$ V, $I_{SS} = 16$ mA and $V_{co} = -4.6$ V; with $v_S = 100$ V, $I_{SS} = 10$ mA and $V_{co} = -3.2$ V; and with $v_S = 50$ V, $I_{SS} = 4$ mA and $V_{co} = -2.1$ V.

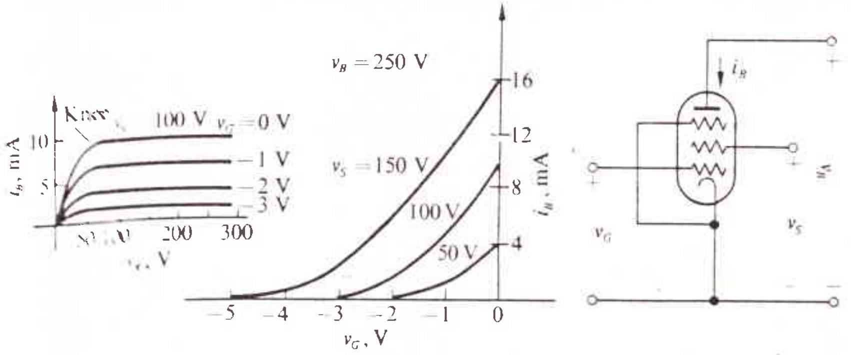


Fig. 4.7–1 Plate current characteristics for a pentode.

An equation such as Eq. (4.7–1) may be derived from theoretical considerations about a device, or it may be matched with the appropriate choice of n, I_{SS}, and V_{co}, to a set of experimentally measured characteristics. It turns out in practice that many field effect transistors have experimentally evaluated characteristics that are best matched by Eq. (4.7–1) with values of n ranging from 1.5 to 2.5. The large majority of them, however, require $n \approx 2$, as we noted previously.

If we now apply a voltage of the form $v_1(t) = V_b + v(t)$, where $v_1(t) > V_{co}$, to a device described by Eq. (4.7–1) or, as a matter of fact, to any device whose output current is a smoothly varying function of its input voltage, then we gain a great deal of insight into the limits of linearity, as well as the nature of the nonlinearity, by expanding i_2 in a Taylor series about the Q-point. Specifically,

$$\begin{aligned} i_2 &= i_2|_{V_b} + \left.\frac{\partial i_2}{\partial v_1}\right|_{V_b} v + \frac{1}{2}\left.\frac{\partial^2 i_2}{\partial v_1^2}\right|_{V_b} v^2 + \frac{1}{2\cdot 3}\left.\frac{\partial^3 i_2}{\partial v_1^3}\right|_{V_b} v^3 + \cdots \\ &= a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \cdots, \end{aligned} \tag{4.7–3}$$

where

$$a_n = \frac{1}{n!}\left.\frac{\partial^n i_2}{\partial v_1^n}\right|_{V_b}.$$

To describe i_2 accurately we must keep all the terms in the Taylor series; however, in practice, if the v_1-i_2 characteristic is smoothly varying, we may require only the first three or four terms to describe the characteristic within a few percent accuracy. Any attempt at further resolution is usually unwarranted because of the lack of precision of the characteristic itself. (As an alternative to evaluating derivatives at the Q-point we could evaluate a_2 and a_3 so that the series gave exact results at two other points spaced about the Q-point. With slowly varying curves the differences between the two methods is small. Since this second approach involves taking the difference between two almost equal numbers, it must be carried out very carefully or the results may be misleading.)

For the device described by Eq. (4.7–1), if we define $V_b = NV_{co}$, where $1 \ge N \ge 0$, then

$$\begin{aligned} a_0 &= I_{SS}(1 - N)^n, \\ a_1 &= -n\frac{I_{SS}}{V_{co}}(1 - N)^{n-1}, \\ a_2 &= \frac{n(n-1)}{2!}\frac{I_{SS}}{V_{co}^2}(1 - N)^{n-2}. \\ &\ \ \vdots \qquad\qquad \vdots \end{aligned} \tag{4.7–4}$$

Clearly, for small ac input voltages v, i_2 is given by

$$i_2 = I_{SS}(1 - N)^n - n\frac{I_{SS}}{V_{co}}(1 - N)^{n-1}v. \tag{4.7–5}$$

Hence we observe that the small-signal transconductance g_m is given by

$$g_m = -n\frac{I_{SS}}{V_{co}}(1 - N)^{n-1}. \tag{4.7–6}$$

Since V_{co} is a negative quantity, we note that g_m is actually positive.

If $v(t) = V_1 \cos \omega t$ such that $v_1 = V_b + V_1 \cos \omega t$, then i_2 (from Eq. 4.7–3) is en by

$$i_2 = a_0 + a_1 V_1 \cos \omega t + a_2 V_1^2 \cos^2 \omega t + a_3 V_1^3 \cos^3 \omega t \tag{4.7–7}$$

ve assume the higher-order terms to be negligible. By using the trigonometric ntities

$$\cos^2 \theta = \frac{1}{2} + \frac{\cos 2\theta}{2} \qquad \text{and} \qquad \cos^3 \theta = \frac{3}{4}\cos \theta + \frac{1}{4}\cos 3\theta,$$

can simplify the expression for i_2 into the Fourier series form

$$i_2 = \underbrace{a_0 + \frac{a_2 V_1^2}{2}}_{I_0} + \underbrace{V_1\left(a_1 + \frac{3a_3 V_1^2}{4}\right)}_{I_1}\cos \omega t + \underbrace{\frac{a_2 V_1^2}{2}}_{I_2}\cos 2\omega t + \underbrace{\frac{a_3 V_1^3}{4}}_{I_3}\cos 3\omega t. \tag{4.7–8}$$

From Eq. (4.7–8) it is now apparent that the relative second-harmonic distortion simply

$$\frac{a_2 V_1}{2\left(a_1 + \dfrac{3a_3 V_1^2}{4}\right)},$$

hich for small values of V_1 reduces to $V_1 a_2/2a_1$. With the aid of Eq. (4.7–4) this can expanded to

$$\frac{V_1 a_2}{2a_1} = \frac{V_1}{4V_{co}} \frac{1-n}{1-N}.$$

$n = \frac{3}{2}$ and $N = \frac{1}{2}$, then for less than 1% second-harmonic distortion $|V_1/V_{co}| \leq 0.04$. /ith $V_{co} = -4$ V, $V_1 \leq 160$ mV.

The large-signal average fundamental transconductance G_m for such a device given by

$$\begin{aligned} G_m = \frac{I_1}{V_1} &= a_1\left(1 + \frac{3V_1^2 a_3}{4a_1}\right) \\ &= -n\frac{I_{SS}}{V_{co}}(1-N)^{n-1}\left[1 + \frac{V_1^2}{8V_{co}^2}\frac{(n-1)(n-2)}{(N-1)^2}\right]. \end{aligned} \tag{4.7–9}$$

he a_3-term has made G_m a function of a signal level. This leads to distortion in arrowband amplifiers. To keep G_m constant within 1%, V_1 must be restricted so hat

$$\frac{V_1^2(n-1)(n-2)}{8V_{co}^2(N-1)^2} < 0.01.$$

Using the same values specified above, we see that $|V_1/V_{co}|$ must be less than 0.283 and V_1 must be less than 1.13 V. Here again a much larger signal may be applied for linear narrowband operation than for linear broadband operation.

If now $v(t) = V_1 \cos \omega t + V_2 \cos \omega_2 t$, then with the aid of Eq. (4.7–3) and several trigonometric identities we can express G_m in the form

$$G_m = \frac{I_1}{V_1} = a_1\left(1 + \frac{3V_1^2 a_3}{4a_1} + \frac{3V_2^2 a_3}{2a_1}\right). \tag{4.7–10}$$

Clearly, if the V_2-term were modulated, then a distorted version of this modulation would appear on the signal output of our narrowband filter centered at ω. This effect of the transference of modulation from one carrier to another is known as cross modulation. One of the advantages of a square-law (or linear) device as a narrowband amplifier is the absence of the cross-modulation effect. (In a true square-law device, $a_3 = 0$.) The first amplifier stage of any receiver should always have a low susceptibility to cross modulation, since it is virtually impossible to have adequate selectivity in front of it to prevent the appearance of some unwanted signals at its input terminals.

4.8 EFFECT OF SERIES RESISTANCE ON THE EXPONENTIAL CHARACTERISTIC

In many practical cases we find that the nonlinearities discussed previously are modified by the inclusion of a resistor in series with the nonlinear element. The general effect of such a resistor is to decrease the transconductance of the composite device (a larger input voltage is required to produce the same level of current that would exist without the resistor) and to "piecewise-linearize" the composite nonlinear characteristic. This linearization provides the means of extending the operating range of a linear amplifier at the expense of its amplification factor. On the other hand, when the composite device is driven with large signals, the principal effect of the series resistor is the increased drive voltage required to produce the required value of input current.

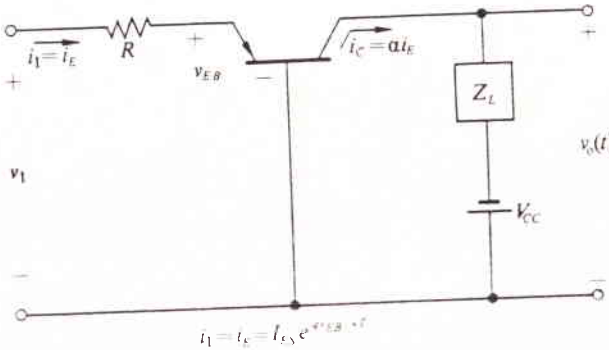


Fig. 4.8–1 A PNP transistor with resistance in series with the emitter.

Figure 4.8–1 shows a transistor with a resistor R in series with the emitter. though such a resistor is sometimes added deliberately, in other instances it is herent in the transistor; for example, it may be the bulk emitter resistance or the trinsic base resistance, which becomes a significant factor when the transistor rries a high current. Thus in most power amplifiers the transistor must be modeled in Fig. 4.8–1.

If we assume that the emitter-base junction of the transistor is described by the ponential characteristic

$$i_E = I_{ES}e^{qv_{EB}/kT}, \tag{4.8–1}$$

en we may relate v_1 to i_1 for the circuit of Fig. 4.8–1 by the expression

$$v_1 = i_1R + \frac{kT}{q}\ln\frac{i_1}{I_{ES}}. \tag{4.8–2}$$

r small-signal operation the incremental input resistance for the circuit takes the rm

$$r'_{\text{in}} = \left.\frac{\partial v_1}{\partial i_1}\right|_{i_1 = I_{\text{dc}}} = R + r_{\text{in}}, \tag{4.8–3}$$

here $r_{\text{in}} = 1/g_{\text{in}} = kT/qI_{\text{dc}}$ and I_{dc} is the dc value of i_1. Consequently, on a small-nal basis we see that the transconductance g'_m given by

$$g'_m = \frac{\alpha}{r'_{\text{in}}} = \frac{\alpha g_{\text{in}}}{1 + g_{\text{in}}R} \tag{4.8–4}$$

creases by a factor of $1 + g_{\text{in}}R$ with increasing R. Specifically, if $\alpha g_{\text{in}} \approx 0.04$ ℧, it would for $I_{\text{dc}} = 1$ mA, then a series resistor of only 10 Ω decreases g'_m to 72% its $R = 0$ value, while a series resistor of 100 Ω decreases g'_m to 20% of its $R = 0$ lue. In the limit as $g_{\text{in}}R$ becomes large compared with unity, g'_m becomes equal to R and is independent of the change in bias current of the transistor.

To observe the "piecewise-linearization" effect of the series resistor on the – i_1 characteristic, we first rearrange Eq. (4.8–2) in the form

$$v_1 - V_0 = i_1R + \frac{kT}{q}\ln\frac{i_1}{I_{\text{dc}}}, \tag{4.8–5}$$

here I_{dc} is the quiescent value of i_1 and

$$V_0 = \frac{kT}{q}\ln\frac{I_{\text{dc}}}{I_{ES}}$$

the quiescent value of v_1 when $i_1 = I_{\text{dc}}$. We then normalize $v_1 - V_0$ to

$$V_{\text{co}} = I_{\text{dc}}(r_{\text{in}} + R) = \frac{kT}{q}(1 + g_{\text{in}}R)$$

and i_1 to I_{dc} to obtain

$$\frac{v_1 - V_0}{V_{\text{co}}} = \left(\frac{g_{\text{in}}R}{1 + g_{\text{in}}R}\right)\frac{i_1}{I_{\text{dc}}} + \frac{\ln(i_1/I_{\text{dc}})}{1 + g_{\text{in}}R}. \tag{4.8–6}$$

The voltage V_{co} may be interpreted as the decrease in v_1 from V_0 which would be required to reduce i_1 to zero (i.e., to cut off the transistor) if the slope of the $v_1 - i_1$ characteristic remained constant at its small-signal value of $1/(r_{\text{in}} + R)$.

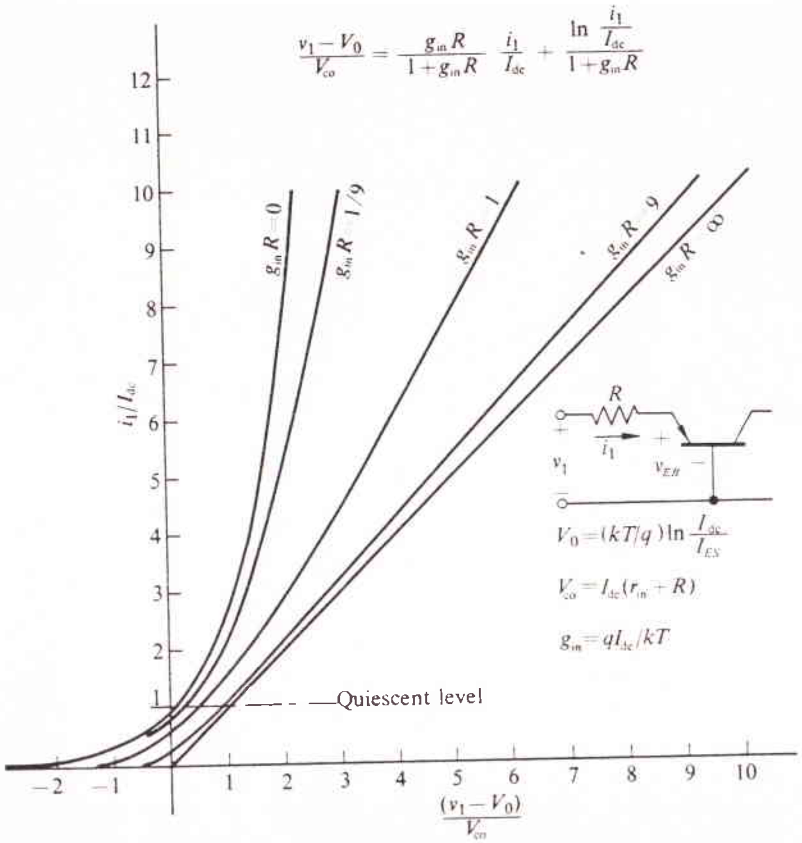


Fig. 4.8–2 Plot of i_1/I_{dc} vs. $v_1 - V_0/V_{\text{co}}$ with $g_{\text{in}}R$ as a parameter.

Finally we plot i_1/I_{dc} vs. $(v_1 - V_0)/V_{\text{co}}$ with $g_{\text{in}}R$ as a parameter, as shown in Fig. 4.8–2. We note from the figure that as $g_{\text{in}}R$ is increased from zero to infinity, the characteristic varies uniformly from the exponential ($g_{\text{in}}R = 0$) characteristic to the two-segment piecewise-linear characteristic ($g_{\text{in}}R = \infty$). In addition, the $g_{\text{in}}R = 1$ characteristic corresponds essentially to the half-way point between the two extremes. Thus for $g_{\text{in}}R \ll 1$ the circuit of Fig. 4.8–1 can be modeled by the transistor characteristic alone, while for $g_{in}R \gg 1$ the circuit can be modeled by the $g_{in}R = \infty$ characteristic, or equivalently,

$$i_1 = \begin{cases} \dfrac{v_1 - V_0}{R}, & v_1 > 0, \\ 0, & v_1 \le 0. \end{cases} \tag{4.8–7}$$

It is apparent that Eq. (4.8–7) describes the circuit of Fig. 4.8–3. This circuit thus provides the model for the circuit of Fig. 4.8–1 when $g_{in}R \gg 1$.

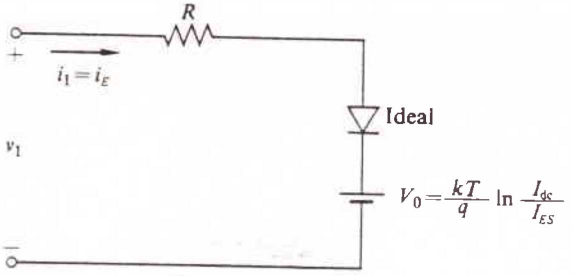


Fig. 4.8–3 Model for the circuit of Fig. 4.8–1 with $g_{in}R \to \infty$.

We also observe from Fig. 4.8–2 that, as $g_{in}R$ is increased from zero, larger linear excursions in i_1 about I_{dc} are possible without gross distortion. In particular, for $g_{in}R \to \infty$, the linear range extends all the way to cutoff ($i_1/I_{dc} = 0$).

If the circuit of Fig. 4.8–1 is to be used as a large-signal narrowband amplifier, the analysis becomes quite difficult for an arbitrary value of $g_{in}R$. Fortunately, an analysis is possible for $g_{in}R = 0$ (which was accomplished in Section 4.5) and for $g_{in}R = \infty$. In addition, the results of the two analyses are sufficiently close to each other that solutions for any value of $g_{in}R$ may be closely approximated.

For the case where $g_{in}R = \infty$, we model the transistor-series resistor combination as shown in Fig. 4.8–4. The current source I_{dc} fixes the average value of the emitter current, while the infinite coupling capacitor ensures that the entire ac voltage reaches the input terminals. (The current source could be placed on the other side of the resistor without affecting the results.) Consequently, for this circuit we observe that, for $V_1 \le I_{dc}R = V_{co}$,

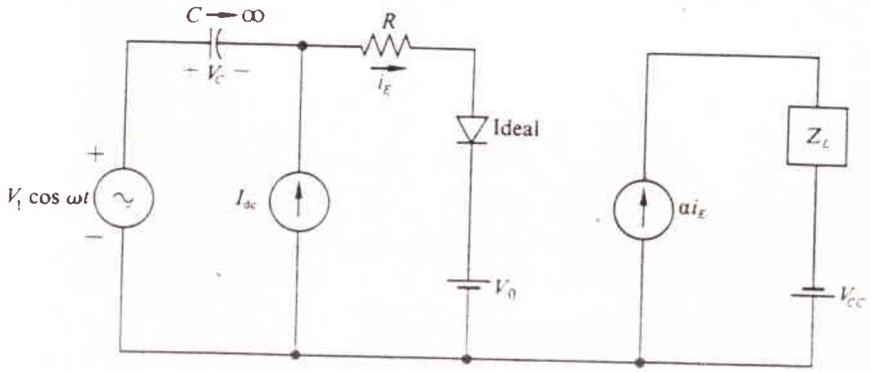


Fig. 4.8–4 Model for large-signal narrowband transistor amplifier with $g_{in}R \to \infty$.

$$i_1 = i_E = I_{dc} + \frac{V_1}{R} \cos \omega t = I_{dc} + I_1 \cos \omega t$$

$$= I_{dc}\left(1 + \frac{V_1}{V_{co}} \cos \omega t\right), \tag{4.8–8}$$

where $I_1 = V_1/R$, or equivalently, $I_1/I_{dc} = V_1/V_{co}$. For $v_1 > V_{co}$ the emitter current takes the form of a periodic train of sine-wave tip pulses, and thus may be written as

$$i_E = I_0 + I_1 \cos \omega t + I_2 \cos 2\omega t + \cdots$$

$$= I_{dc}\left(1 + \frac{I_1}{I_0} \cos \omega t + \frac{I_2}{I_0} \cos 2\omega t + \cdots\right), \tag{4.8–9}$$

where (from the appendix to this chapter)

$$I_0 = \frac{V_1 - V_x}{\pi R} \frac{\sin \phi - \phi \cos \phi}{1 - \cos \phi},$$

$$I_1 = \frac{V_1 - V_x}{\pi R} \frac{\phi - \cos \phi \sin \phi}{1 - \cos \phi},$$

$$I_n = \frac{2(V_1 - V_x)}{\pi R} \frac{\cos \phi \sin n\phi - n \sin \phi \cos n\phi}{n(n^2 - 1)(1 - \cos \phi)}, \qquad n > 2.$$

$$\phi = \cos^{-1} \frac{V_x}{V_1}, \qquad V_x = -V_C - V_0,$$

and V_C is the developed dc capacitor voltage. By noting that

$$V_1 - V_x = V_1(1 - \cos \phi) \qquad \text{and} \qquad I_0 = I_{dc},$$

we can relate V_1/V_{co} directly to ϕ by the equation

$$\frac{V_1}{V_{co}} = \frac{\pi}{\sin \phi - \phi \cos \phi}. \tag{4.8–10}$$

In addition, I_1/I_0 can be related to ϕ in the form

$$\frac{I_1}{I_0} = \frac{\phi - \sin \phi \cos \phi}{\sin \phi - \phi \cos \phi}, \tag{4.8–11}$$

and in a similar fashion

$$\frac{I_n}{I_0} = \frac{2(\cos \phi \sin n\phi - n \sin \phi \cos n\phi)}{n(n^2 - 1)(\sin \phi - \phi \cos \phi)}. \qquad n \ge 2. \tag{4.8–12}$$

Consequently, for any value of ϕ, a value of V_1/V_{co} and I_n/I_{dc} may be determined from which a unique relationship between V_1/V_{co} and I_n/I_{dc} can be established. Table 4.8–1 contains values of $I_1/I_0 = I_1/I_{dc}$ vs. V_1/V_{co} which are plotted in Fig. 4.8–5 with the designation $g_{in}R = \infty$. On the same set of coordinates is plotted the $g_{in}R = 0$ curve, which, of course, is just the top curve of Fig. 4.5–5, since with $R = 0$, $I_1/I_0 = 2I_1(x)/I_0(x)$ and $V_1/V_{co} = x$. We observe that the two limiting curves are almost identical; hence the ratio of I_1/I_0 for any value of $g_{in}R$ can be closely estimated from Fig. 4.8–5.† For example, if $g_{in}R = 1$ and $V_1/V_{co} = 10$, it is reasonable to approximate I_1/I_0 as 1.95, which is half-way between the $g_{in}R = 0$ and the $g_{in}R = \infty$ curves. If, in addition, $I_{dc} = 2$ mA, then $I_1 = 3.7$ mA, from which the output voltage may be written as $-V_{CC} + I_1R_L \cos \omega t$ for the case where Z_L is a high-Q tuned circuit which is resonant at ω with a shunt resistance of R_L.

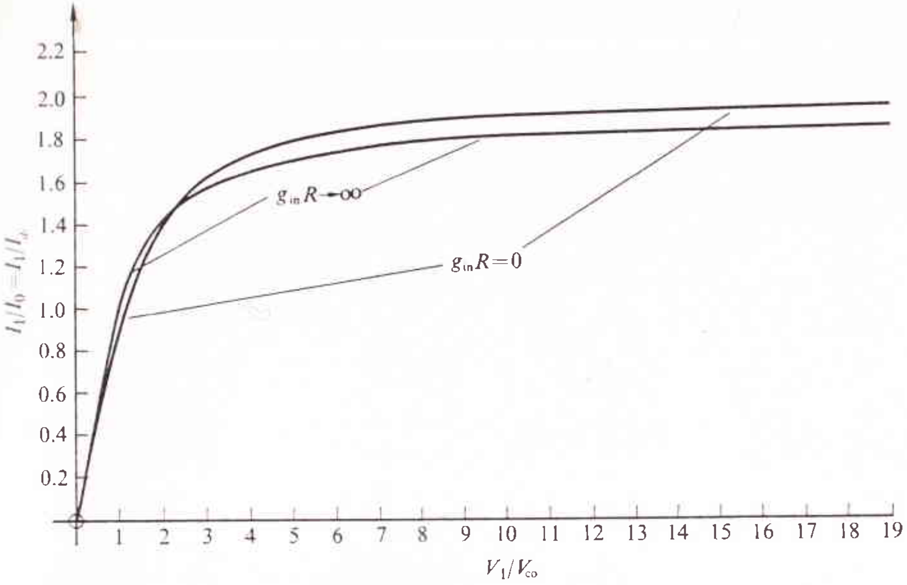


Fig.4.8–5 Plot of I_1/I_{dc} vs. V_1/V_{co}.

Table 4.8–1

ϕ	$I_1/I_0 = I_1/I_{dc}$	V_1/V_{co}
40°	1.88	28.8
60°	1.80	9.17
80°	1.65	4.25
100°	1.49	2.44
120°	1.32	1.64
140°	1.17	1.25
160°	1.05	1.06
180°	1	1

For the case where the circuit of Fig. 4.8–1 is driven by a sinusoidal voltage of amplitude V_1 and followed by a tuned circuit resonant at the frequency of the input sinusoid, it is again convenient to define a large-signal average fundamental transconductance as

$$G_m = \frac{\alpha I_1}{V_1} = \frac{I_1/I_{dc}}{V_1/V_{co}} \frac{\alpha I_{dc}}{V_{co}} = \frac{I_1/I_{dc}}{V_1/V_{co}} g'_m, \tag{4.8–13}$$

where g'_m is the small-signal transconductance evaluated at the Q-point. Clearly, values for

$$\frac{G_m}{g'_m} = \frac{I_1/I_{dc}}{V_1/V_{co}} \tag{4.8–14}$$

vs. V_1/V_{co} may be determined directly from Fig. 4.8–5.

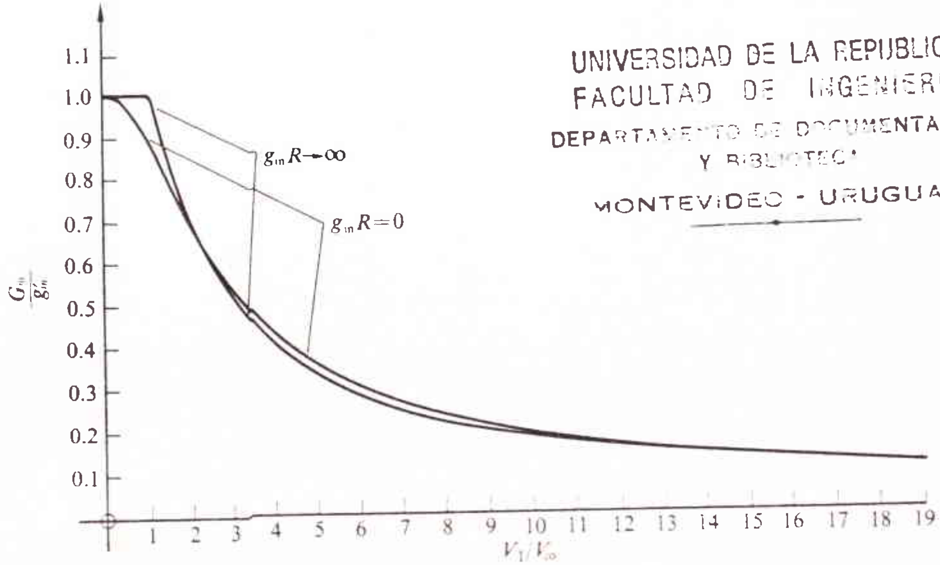


Fig. 4.8–6 Plot of G_m/g'_m vs. V_1/V_{co}.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION
Y BIBLIOTECA
MONTEVIDEO - URUGUAY

---

† It is not unreasonable for the I_1/I_{dc} vs. V_1/V_{co} curves to form a tight set with $g_{in}R$ as a parameter, since all the curves have the same asymptotes for both small and large values of V_1/V_{co}. Clearly, for $V_1/V_{co} \ll 1$, $I_1/I_{dc} = V_1/V_{co}$. Therefore, regardless of the value of $g_{in}R$, the curves approach the origin in an identical fashion. In addition, for large values of V_1/V_{co}, if the average current is to remain constant at I_{dc}, the transistor must remain cut off for most of the cycle. Consequently, the current flows in narrow pulses and $I_1/I_{dc} \to 2$. Again, all the curves approach infinity in the same fashion.

A plot of G_m/g'_m vs. V_1/V_{co} for $g_{in}R = 0$ and $g_{in}R = \infty$ is shown in Fig. 4.8–6. Here again we see that the curves are sufficiently close to permit an accurate estimate of G_m/g'_m for any value of $g_{in}R$. It is now quite apparent that the ratio of I_1/I_0 is determined primarily by the ratio of V_1/V_{co} and is almost independent of $g_{in}R$. However, the value of

$$V_{co} = I_{dc}(r_{in} + R) = \frac{kT}{q}(1 + g_{in}R)$$

increases with $g_{in}R$. Consequently, as $g_{in}R$ is increased, the drive voltage must be increased by a factor of $1 + g_{in}R$ to maintain the same fundamental output current.

Example 4.8–1 For the circuit shown in Fig. 4.8–7 determine an expression for $v_o(t)$.

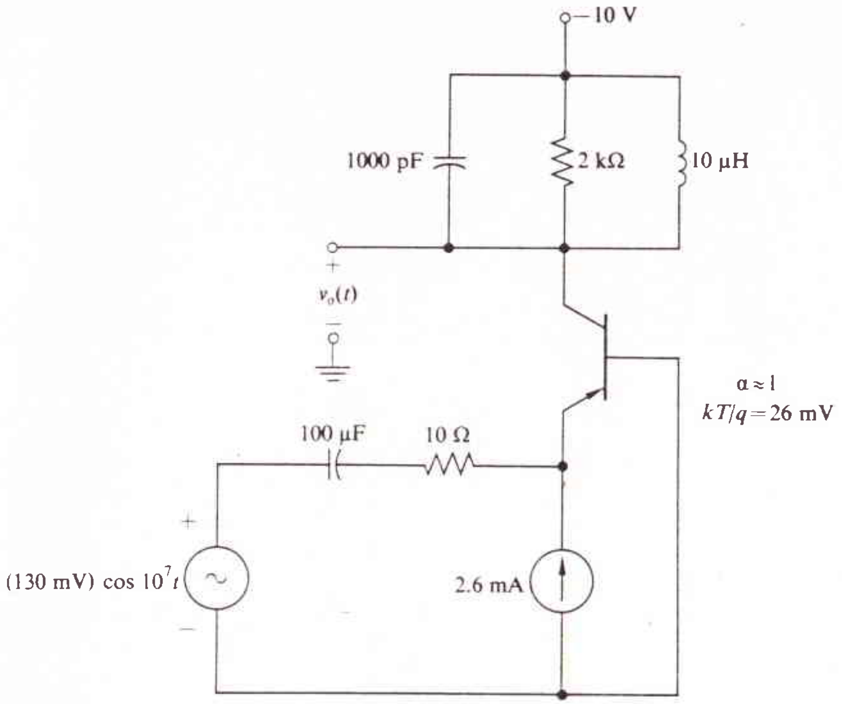


Figure 4.8–7

Solution. Since the tuned circuit in the collector has $Q_T = 20$ and is tuned to $\omega = 10^7$ rad/sec, we note that

$$v_o(t) = -10\text{ V} + \alpha(2\text{ k}\Omega)I_1 \cos 10^7t,$$

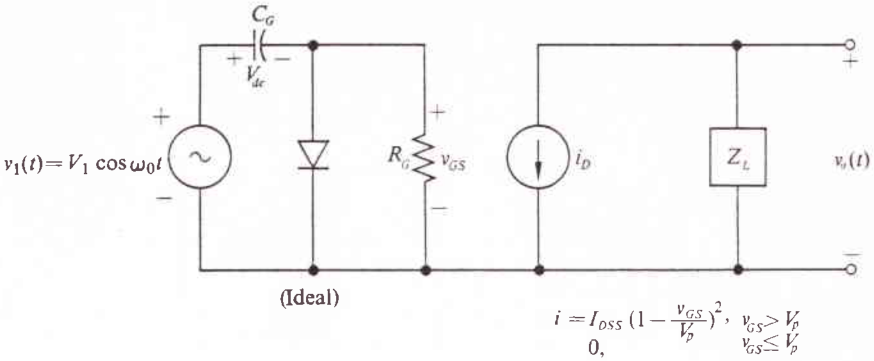
where I_1 is the fundamental component of the emitter current. We obtain I_1 by first noting that, with an average emitter current of 2.6 mA, $g_{in} = 1/(10\ \Omega)$ and therefore $g_{in}R = 1$. In addition, $V_{co} = (kT/q)(2) = 52$ mV and thus $V_1/V_{co} = 2.5$. Consequently, from Fig. 4.8–5 we estimate $I_1/I_{dc} \approx 1.5$; thus $I_1 = 3.9$ mA and

$$v_o(t) = (-10\text{ V}) + (7.8\text{ V}) \cos 10^7t.$$

4.9 CLAMP-BIASED SQUARE-LAW DEVICE

In many large-signal junction FET circuits the bias is established by clamping the positive peaks of the gate to source voltage to zero (or the turn-on voltage of the gate to source diode). In particular, most FETs employed in self-limiting oscillators are biased in this fashion; hence we shall analyze the large-signal characteristics for this type of circuit in this section.

We begin our analysis with the clamp-biased square-law amplifier model shown in Fig. 4.9–1, which is a close first-order approximation to the actual junction FET amplifiers shown in Fig. 4.9–2. Clearly, in this model, if the R_GC_G time constant is



C_G
+
−
V_{dc}
+
$v_1(t) = V_1 \cos \omega_0 t$
−
(Ideal)
+
R_G
v_{GS}
−
i_D
Z_L
+
$v_o(t)$
−
$i = I_{DSS}\left(1 - \frac{v_{GS}}{V_p}\right)^2, \quad v_{GS} > V_p$
$0, \quad v_{GS} \le V_p$

Fig. 4.9–1 Large-signal clamp-biased square-law tuned amplifier.

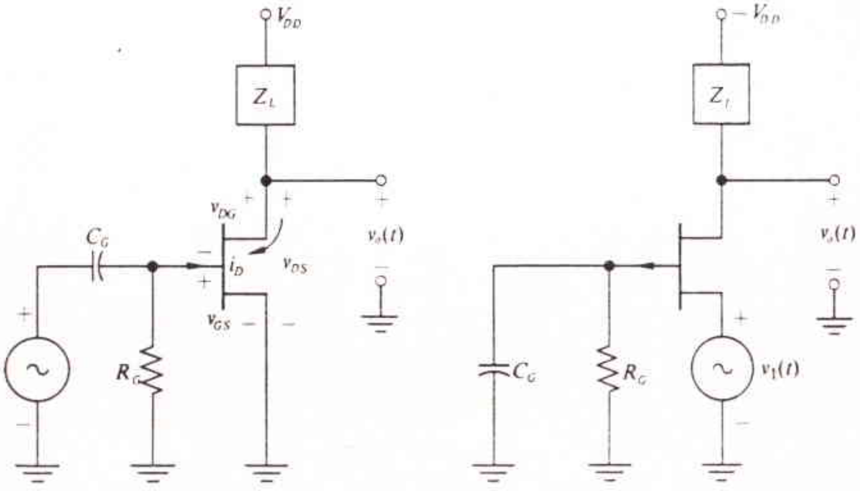


Fig. 4.9–2 Clamp-biased junction FET circuits.

much greater than $T = 2\pi/\omega$, the capacitor voltage charges to the peak value of $v_1(t)$ through the diode and remains constant at that value; hence

$$V_{dc} = V_1 \tag{4.9–1}$$

and

$$v_{GS} = V_1(\cos \omega_0 t - 1). \tag{4.9–2}$$

Thus we observe that the input voltage v_{GS} has its peaks clamped to zero. Consequently $i_D(t)$ has the form shown in Fig. 4.4–3, with $I_p = I_{DSS}$, $V_x = V_1 + V_p$, and $V_x/V_1 = 1 + V_p/V_1$. Therefore, with the aid of Eq. (4.4–10) (for $V_x/V_1 \leq -1$ or $V_1 \leq V_p/2$) and Fig. 4.4–5 (for $-1 < V_x/V_1 < 1$ or $V_1 > V_p/2$) we can plot I_0/I_p, I_1/I_p, and I_2/I_p vs. $-V_1/V_p$ as shown in Fig. 4.9–3, where I_0, I_1, and I_2 are the average, the fundamental, and the second-harmonic components of i_D respectively.

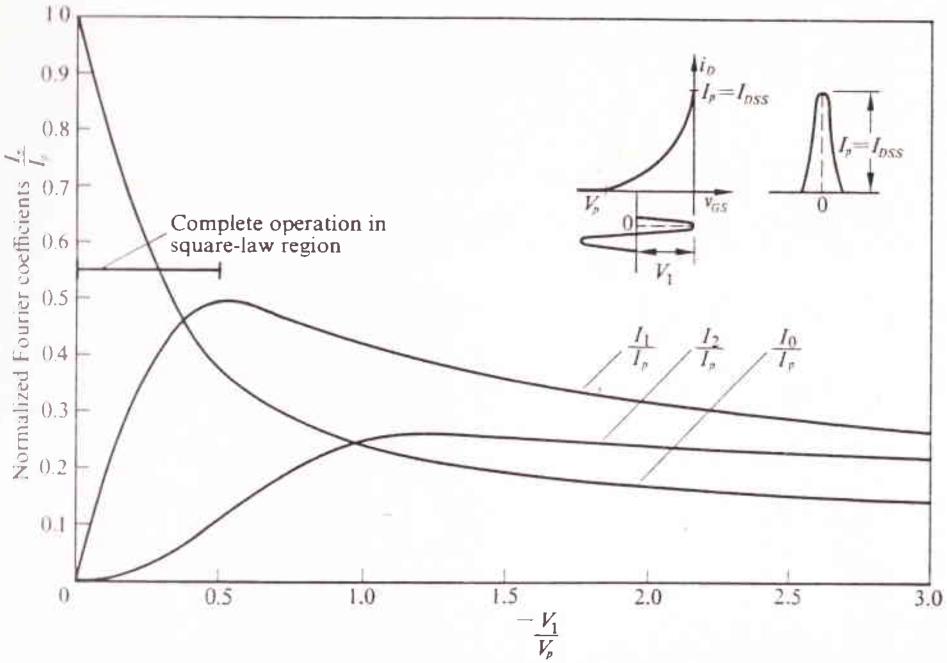


Fig. 4.9–3 Plot of I_n/I_p vs. V_1/V_p for a peak-clamped sine-wave drive to a half-square-law amplifier ($n = 0, 1, 2$).

If Z_L is a high-Q_T parallel RLC circuit tuned to the fundamental of $i_D(t)$ with a resistance R_L, the output voltage $v_o(t)$ is given by

$$v_o(t) = V_{DD} - R_L I_{DSS} \frac{I_1}{I_p} \cos \omega_0 t. \tag{4.9–3}$$

where I_1/I_p can be determined from Fig. 4.9–3 once $-V_1/V_p$ is specified. For example, if $I_{DSS} = 4$ mA, $V_p = -4$ V, $V_1 = 2$ V, $V_{DD} = 10$ V, and $R_L = 2$ kΩ, then $-V_1/V_p = \frac{1}{2}$ and in turn $I_1/I_p = I_1/I_{DSS} = 0.5$; hence

$$v_o(t) = (10 \text{ V}) - (4 \text{ V}) \cos \omega t.$$

In this example the minimum value of $v_o = v_{DS}$ is 6 V; since this is greater than $-V_p$, we can be sure that the FET remains within its saturation region, for which the square-law model of Fig. 4.9–1 is valid. (In general, an FET remains within its saturation region if $v_{DG} > -V_p$, or equivalently, $v_{DS} - v_{GS} > -V_p$ for all time.)

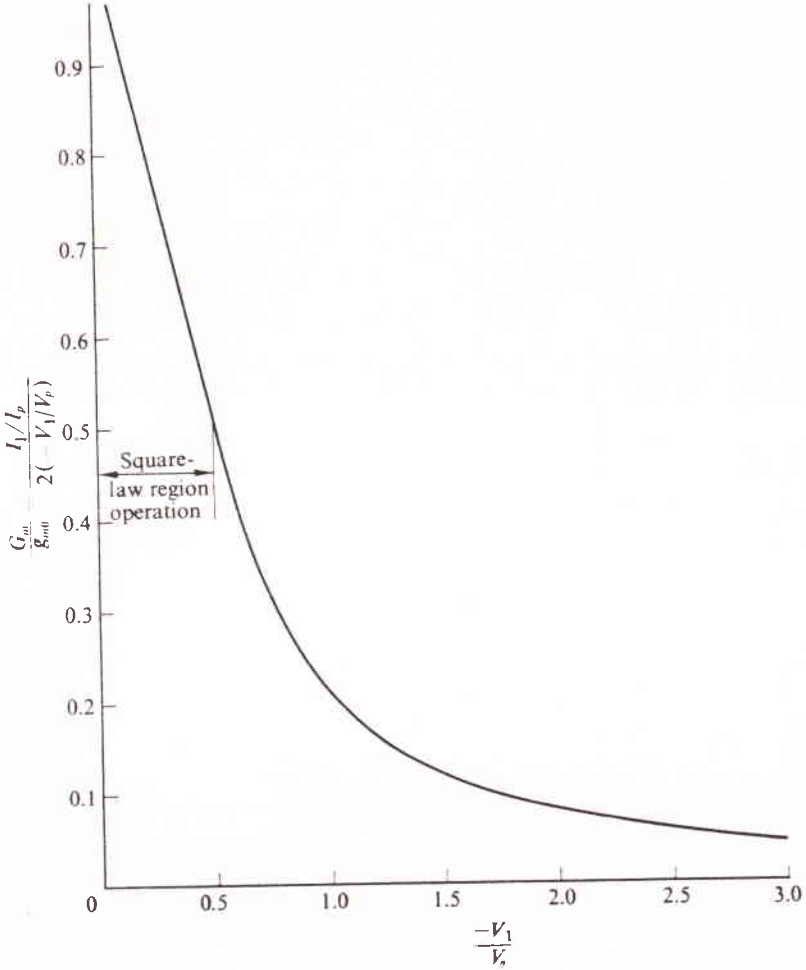


Fig. 4.9–4 Plot of $\frac{G_m}{g_{m0}} = \frac{I_1/I_p}{2(-V_1/V_p)}$ for a peak-clamped sine-wave drive to a half-square-law amplifier.

For the clamp-biased model of Fig. 4.9–1 we may also define a large-signal average fundamental transconductance

$$G_m = \frac{I_1}{V_1} = \frac{2I_{DSS}(I_1/I_p)}{-V_p(2)(-V_1/V_p)} = g_{mo}\frac{I_1/I_p}{2(-V_1/V_p)}, \tag{4.9–4}$$

where $g_{mo} = 2I_{DSS}/-V_p$ is the small-signal transconductance evaluated at $v_{GS} = 0$ (cf. Eq. 4.4–8). A plot of

$$\frac{G_m}{g_{mo}} = \frac{I_1/I_p}{2(-V_1/V_p)}$$

vs. V_1/V_p obtained directly from Fig. 4.9–3 is shown in Fig. 4.9–4. With G_m obtained from Fig. 4.9–4, $v_o(t)$ may be written in the equivalent form

$$v_o = V_{DD} - G_m R_L V_1 \cos \omega_0 t. \tag{4.9–5}$$

The effect of a 0.7 V diode turn-on voltage can be included in the previous results by defining the drain current at this turn-on voltage as I^*_{DSS} and substituting appropriately.

Effect of Drain Resistance

In the analysis to this point, we have assumed that i_D is a half-square-law function of v_{GS} and is independent of $v_o = v_{DS}$. This assumption is justified only if R_L is small compared with the effective output resistance of the square-law device. If R_L is not small, a more exact model for i_D must be employed.

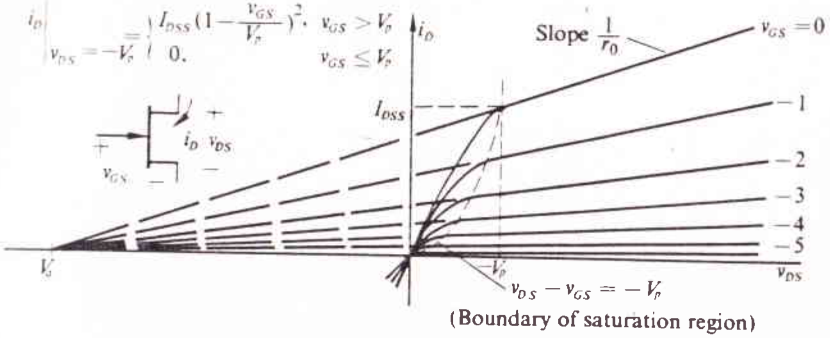


Fig. 4.9–5 Typical set of FET drain characteristics.

Figure 4.9–5 illustrates a typical set of drain characteristics for an *N*-channel junction FET which clearly indicates that no characteristic for constant v_{GS} is horizontal (a condition which would have to be satisfied for i_D to be independent of v_{DS}). On the contrary, each constant-v_{GS} characteristic has a slope which, when extrapolated, causes all of the constant-v_{GS} characteristics to meet at a common point V_x on the negative v_{DS}-axis. Thus, if for $v_{DS} = -V_p$,

$$i_D = I_{DSS}\left(1 - \frac{v_{GS}}{V_p}\right)^2, \qquad v_{GS} > V_p, \tag{4.9–6}$$

then for any other value of v_{DS} we obtain from the geometry of Fig. 4.9–5

$$i_D = I_{DSS}\left(1 + \frac{v_{DS} + V_p}{r_0 I_{DSS}}\right)\left(1 - \frac{v_{GS}}{V_p}\right)^2, \tag{4.9–7}$$

where $1/r_0$ is the slope of the $v_{GS} = 0$ characteristic. Equation (4.9–7) includes the effect of v_{DS} on i_D.

If we now employ the expression for i_D given by Eq. (4.9–7) in the model of Fig. 4.9–1, in which Z_L is a high-Q_T parallel *RLC* circuit tuned to the fundamental frequency of i_D and having a resistance R_L, we obtain

$$\begin{aligned} i_D = I_{DSS}&\left(1 + \frac{V_{DD} + V_p}{r_0 I_{DSS}} - \frac{I_{D1} R_L \cos \omega_0 t}{r_0 I_{DSS}}\right) \\ &\times \left(\frac{I_0}{I_p} + \frac{I_1}{I_p}\cos \omega_0 t + \frac{I_2}{I_p}\cos 2\omega_0 t + \cdots\right), \end{aligned} \tag{4.9–8}$$

where I_0/I_p, I_1/I_p, and I_2/I_p are the coefficients presented in Fig. 4.9–3, and I_{D1} is the amplitude of the fundamental component of i_D. The fundamental component of i_D can also be obtained from Eq. (4.9–8) in the form

$$I_{D1} = I_{DSS}\left(1 + \frac{V_{DD} + V_p}{r_0 I_{DSS}}\right)\frac{I_1}{I_p} - \frac{I_{D1} R_L}{r_0}\left(\frac{I_0}{I_p} + \frac{I_2}{2I_p}\right). \tag{4.9–9}$$

Hence Eq. (4.9–9) may be solved for I_{D1} to obtain

$$\begin{aligned} I_{D1} &= \frac{I_{DSS}\left(1 + \dfrac{V_{DD} + V_p}{r_0 I_{DSS}}\right)\dfrac{I_1}{I_p}}{1 + \dfrac{R_L}{r_0}\left(\dfrac{I_0}{I_p} + \dfrac{I_2}{2I_p}\right)} \\ &= I'_{DSS}\frac{\dfrac{I_1}{I_p}}{1 + \dfrac{R_L}{r_0}\left(\dfrac{I_0}{I_p} + \dfrac{I_2}{2I_p}\right)}, \end{aligned} \tag{4.9–10}$$

where

$$I'_{DSS} = I_{DSS}\left(1 + \frac{V_{DD} + V_p}{r_0 I_{DSS}}\right)$$

is the value of i_D obtained with $v_{GS} = 0$ and $v_{DS} = V_{DD}$.

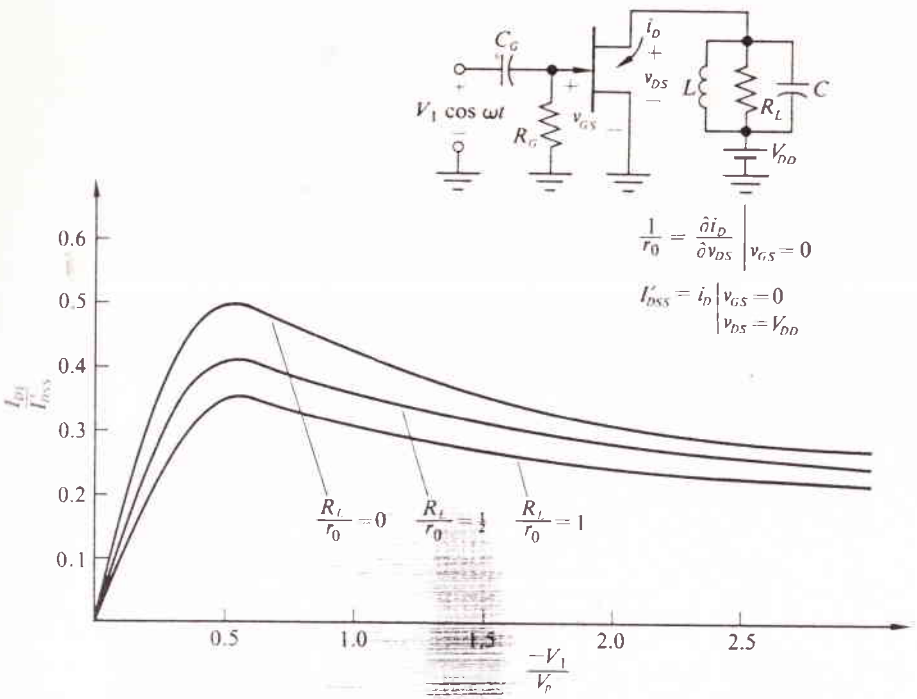


Fig. 4.9–6 Plot of I_{D1}/I_{DSS} vs. $-V_1/V_p$ with R_L/r_0 as a parameter.

Figure 4.9–6 includes a plot of

$$\frac{\dfrac{I_1}{I_p}}{1 + \dfrac{R_L}{r_0}\left(\dfrac{I_0}{I_p} + \dfrac{I_2}{2I_p}\right)}$$

for values of R_L/r_0 of 1, $\frac{1}{2}$, and 0. The curve for $R_L/r_0 = 0$ corresponds to the I_1/I_p curve of Fig. 4.9–3.

As an example of using the curves of Fig. 4.9–6, we consider the circuit shown in Fig. 4.9–7. We observe for this circuit that

$$I'_{DSS} = (2\text{ mA})\left(1 + \frac{16\text{ V}}{40\text{ V}}\right) = 2.8\text{ mA}, \qquad \frac{R_L}{r_0} = \frac{1}{2}, \qquad \text{and} \qquad -\frac{V_1}{V_p} = \frac{1}{2};$$

hence from Fig. 4.9–6 we obtain

$$I_{D1} = (2.8)(0.41)\text{ mA} = 1.15\text{ mA}.$$

In addition, since the resonant frequency of the drain circuit is 10^7 rad/sec and the circuit $Q_T = 10$.

$$v_o = (20\text{ V}) - (11.5\text{ V})\cos 10^7 t.$$

Since $v_o(t) > -V_p$, the FET remains within its saturation region and the assumed FET model is valid.

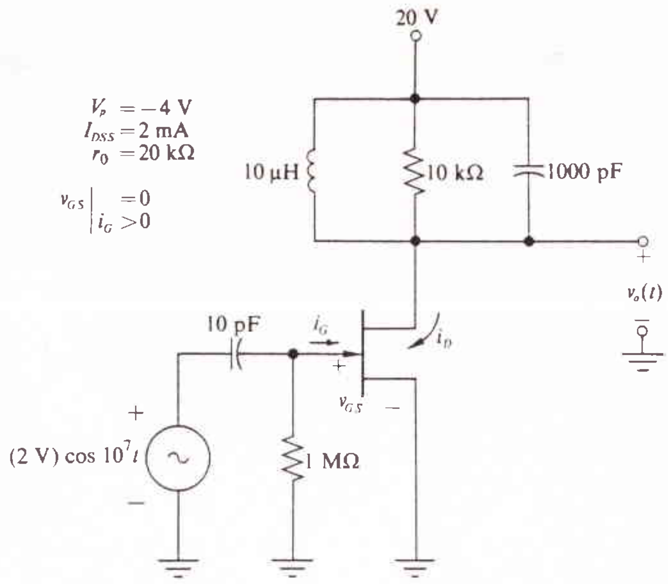


Fig. 4.9–7 Tuned FET amplifier.

PROBLEMS

4.1 Assume that $v_i(t) = 3$ V $\sin \omega_0 t$ in Fig. 4.P–1. Determine the dc, fundamental, and second-harmonic components of $i_2(t)$ for values of V_{dc} of 0, ±2 V, and ±4 V. Sketch the output waveshape for each case.

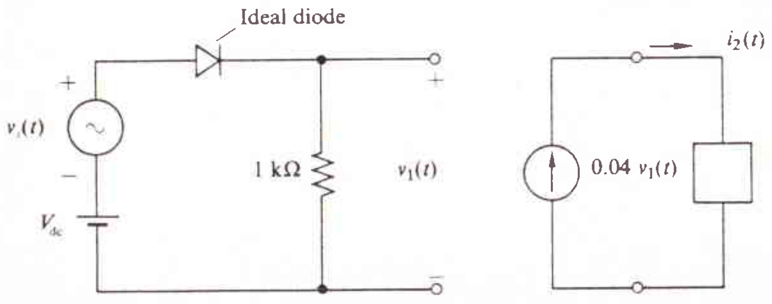


Figure 4.P–1

4.2 Assume that the diodes in Fig. 4.P–2 are ideal. Find the dc, fundamental, second-harmonic, and third-harmonic components of $i_{2\,\text{k}\Omega}(t)$ in this circuit.

4.3 How would the results of Problem 4.2 be changed if the magnitude of the battery in the lower branch were increased to 3 V?

4.4 Assuming that the diodes in Fig. 4.P–3 are ideal, sketch the waveshape of $v_o(t)$ and calculate its third-harmonic component.

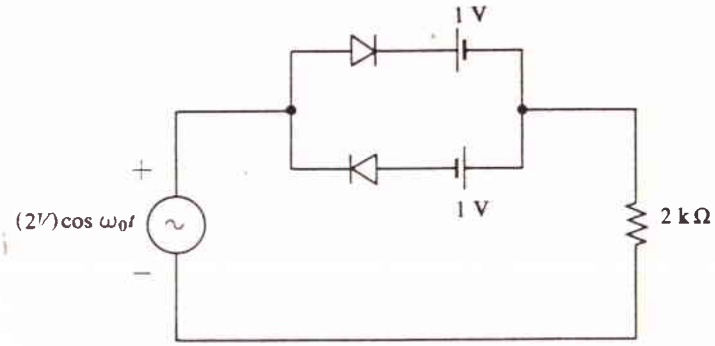


Figure 4.P–2

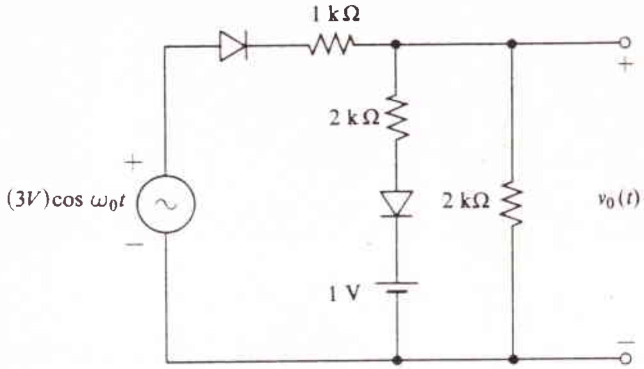


Figure 4.P–3

4.5 Assuming ideal diodes in Fig. 4.P–4, sketch $v_o(t)$ and determine the dc, fundamental, second-harmonic, and third-harmonic components.

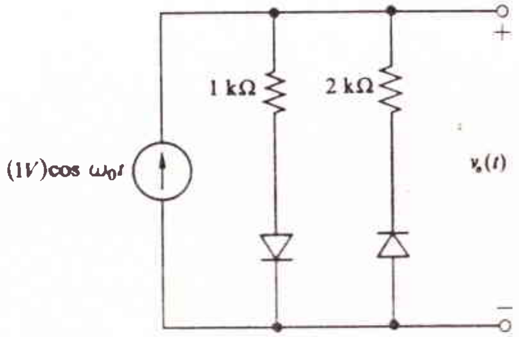


Figure 4.P–4

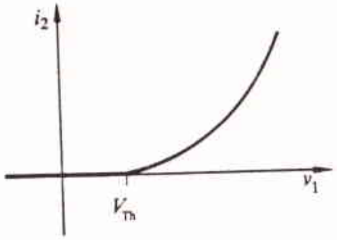


Figure 4.P–5

4.6 Assume a device with the characteristic shown in Fig. 4.P–5. For this device,

$$i_2 = \begin{cases} (v_1 - V_{Th})^2\dfrac{\beta}{2}, & v_1 > V_{Th}, \\ 0, & v_1 \le V_{Th}, \end{cases}$$

where $V_{Th} = +2$ V and $\beta = 400\ \mu\text{A/V}^2$. Evaluate the large-signal average transconductance for this device for the case where

$$v_1 = V_{dc} + V_1 \cos\omega t \qquad \text{and} \qquad V_{dc} > 2\text{ V}$$

while

$$V_1 \le |V_{dc} - V_{Th}|.$$

4.7 Assuming that $V_{dc} = 0$ and $V_1 = 3\cos\omega t$, find the dc, fundamental, second-harmonic, and third-harmonic components of i_2 in the device of Problem 4.6.

4.8 Sketch the output current waveshape for a device with an exponential current-voltage relationship

$$i_2 = 1 \times 10^{-10} e^{v_1/0.026}\text{ A} \qquad \text{when} \qquad v_1 = 0.520\text{ V}\cos\omega t.$$

Specify the "conduction angle." Find the dc, fundamental, second-harmonic, and third-harmonic components of this current. What is the large-signal average transconductance for this device with this driving waveshape?

4.9 In the circuit of Fig. 4.P–6, the diode conducts the same current as the emitter of the transistor when $v_{BE} = v_{\text{diode}}$. If $I_{dc} = 2$ mA, $V_{CC} = +6$ V, and Z_L is a 2.7 kΩ resistor, how large can V_1 be before there is 2% distortion in the output voltage $v_o(t)$? If Z_L is replaced by a parallel tuned circuit resonant at ω_0 and with a resistive term of 2.7 kΩ, how large can V_1 be before there is a 2% departure in linearity between the peak amplitudes of input and output ac voltages? (*Hint:* Note the similarity of this circuit to the differential pair of transistors.)

4.10 In the circuit of Fig. 4.P–6, assume that a 50 Ω resistor is placed both in series with the diode and directly in series with the emitter. Plot a curve of collector current vs. base-ground voltage over the range $1.95 > i_C > 0.05$ mA. Compare this curve with the appropriate data from Fig. 4.6–2 and explain the significance of the resistors in extending the signal-handling capacity of the amplifier. What are the small-signal voltage gains ($R_L = 2.7$ kΩ) for the two different cases ($R_E = 50\ \Omega$ and $R_E = 0$)? State and explain assumptions.

4.11 Determine the output-tuned circuit voltage in Fig. 4.P–7. The tuned circuit is resonant at ω_0 and has $Q_T = 20$.

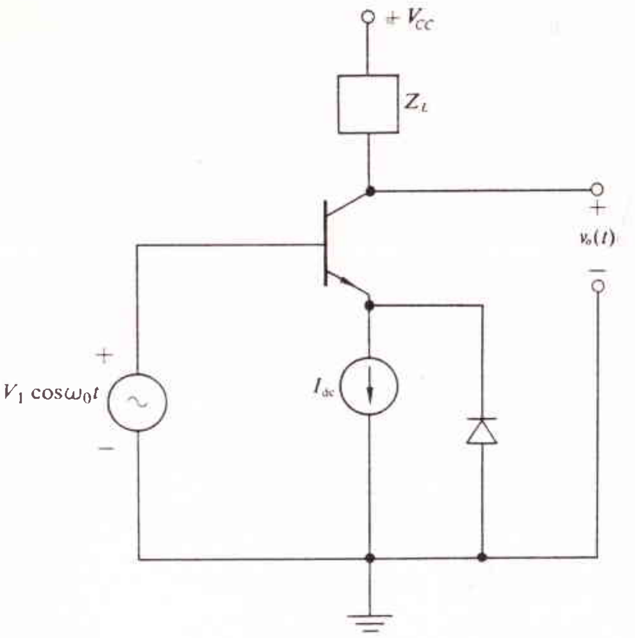


Figure 4.P–6

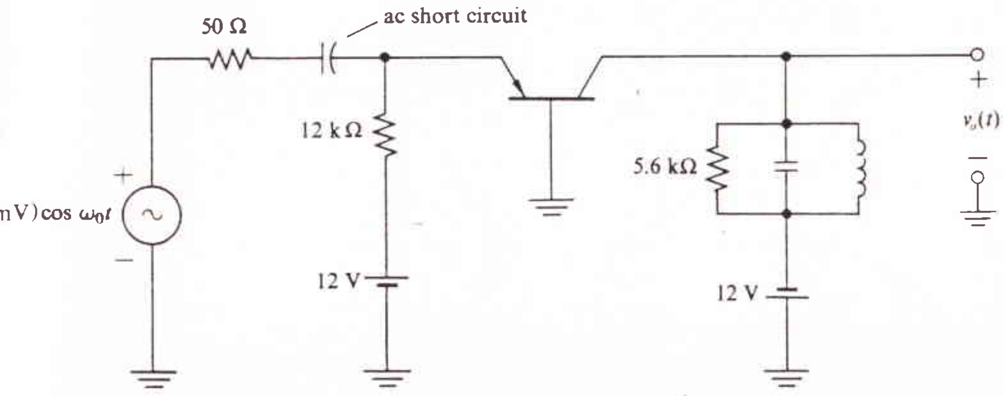


Figure 4.P–7

4.12 The FET in Fig. 4.P–8 is silicon and has an I_{DSS} magnitude of 6 mA and a V_p magnitude of 4 V. What value of V_1 will make the output dc current equal to 1.8 mA? What value of ac voltage will result in this case? What is the circuit's voltage "gain"? How much power is being dissipated in the FET during this operation? The tuned circuit is resonant at 10^7 rps.

4.13 a) For the circuit of Fig. 4.P–9, determine an expression for $v_o(t)$ assuming a high value of Q_T.

b) Determine the minimum value of Q_T which keeps the total harmonic distortion below 1%.

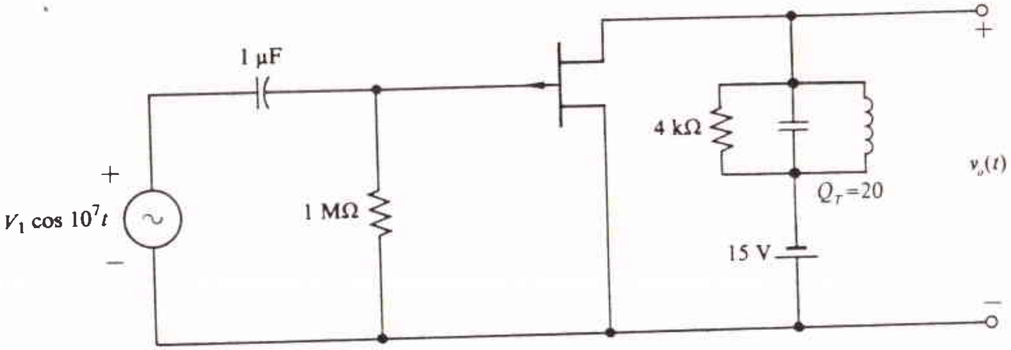


Figure 4.P–8

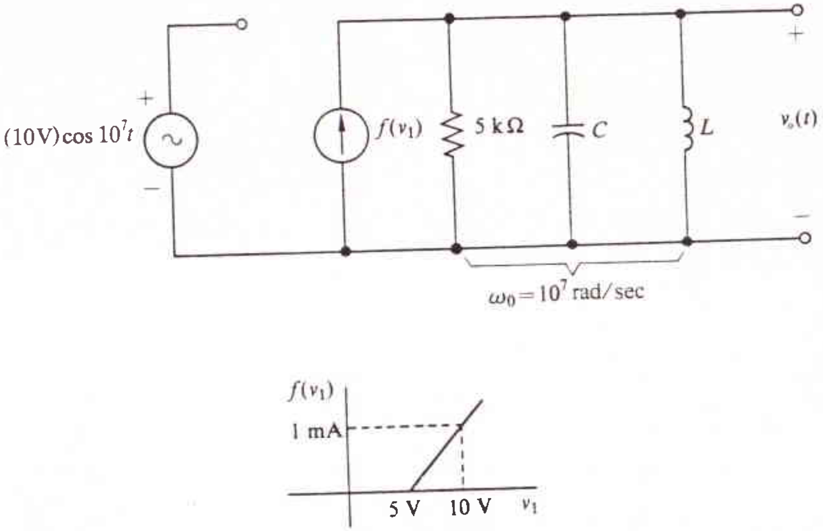


Figure 4.P–9

4.14 For each of the circuits in Fig. 4.P–10, determine and sketch $v_o(t)$ and $i_E(t)$. In each case $v_i(t) = (260\text{ mV}) \cos 10^7 t$.

4.15 For the circuit shown in Fig. 4.P–11, determine $v_o(t)$ where $L = 5\ \mu\text{H}$, $C = 2000$ pF, and $v_i(t) = (5\text{ V}) \cos 10^7 t$.

4.16 Repeat Problem 4.15 for $L = 3.3\ \mu\text{H}$ and $C = 330$ pF.

4.17 Repeat Problem 4.15 with D_1 removed and with $L = 10\ \mu\text{H}$ and $C = 1000$ pF.

4.18 a) For the circuit of Fig. 4.P–12, determine the small-signal voltage amplification $A_V = v_o/v_i$.

b) Make a plot of v_o vs. v_i as v_i varies from -8 V to $+8$ V. What is the effect of the emitter resistors?

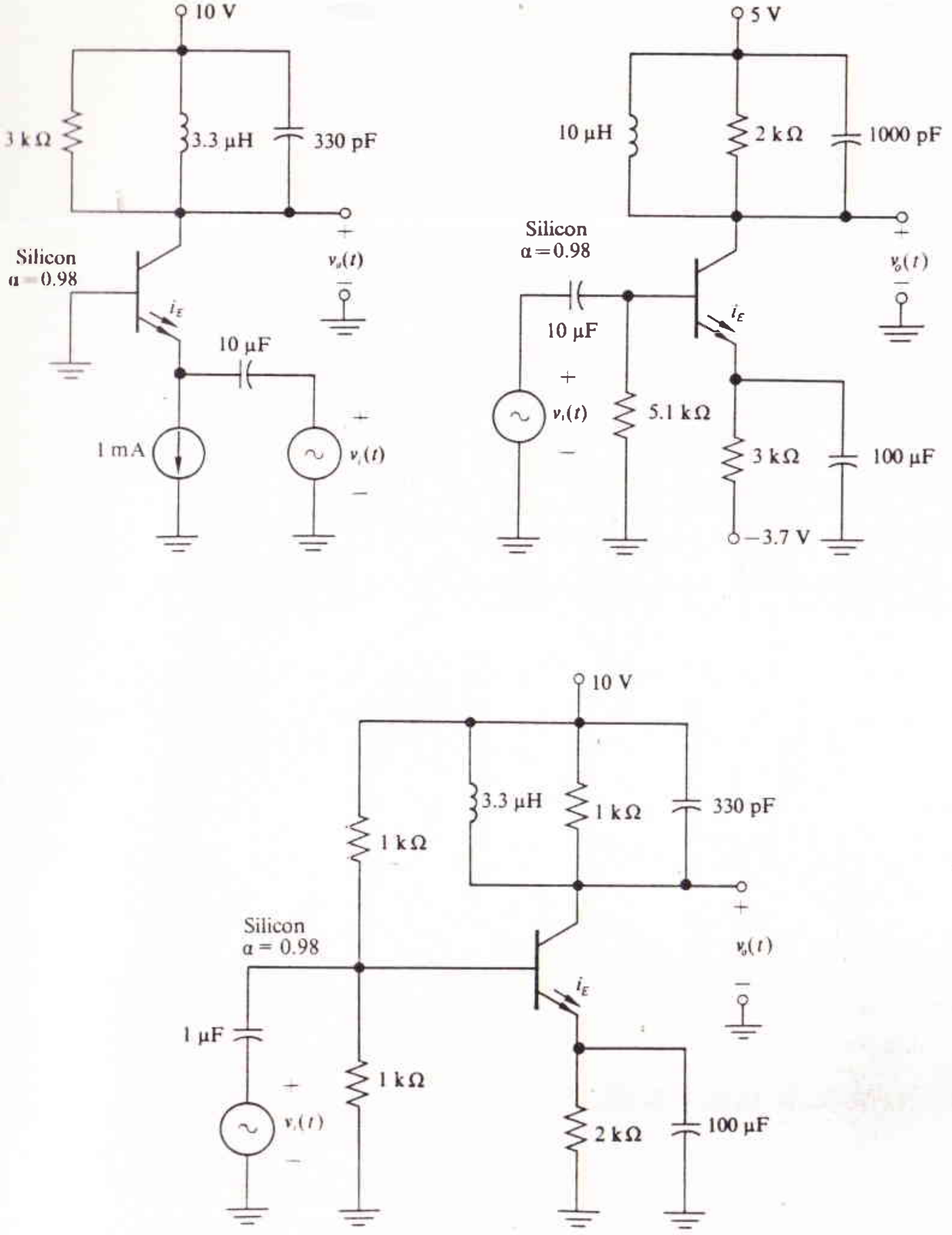


Figure 4.P–10

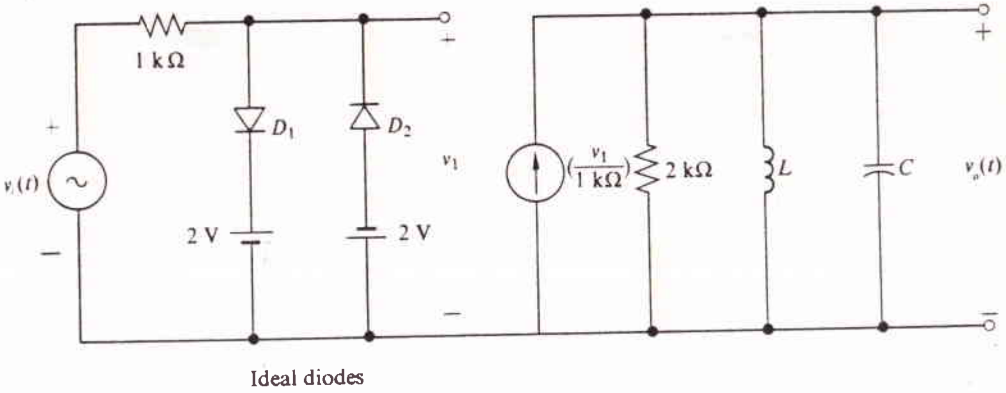


Figure 4.P–11

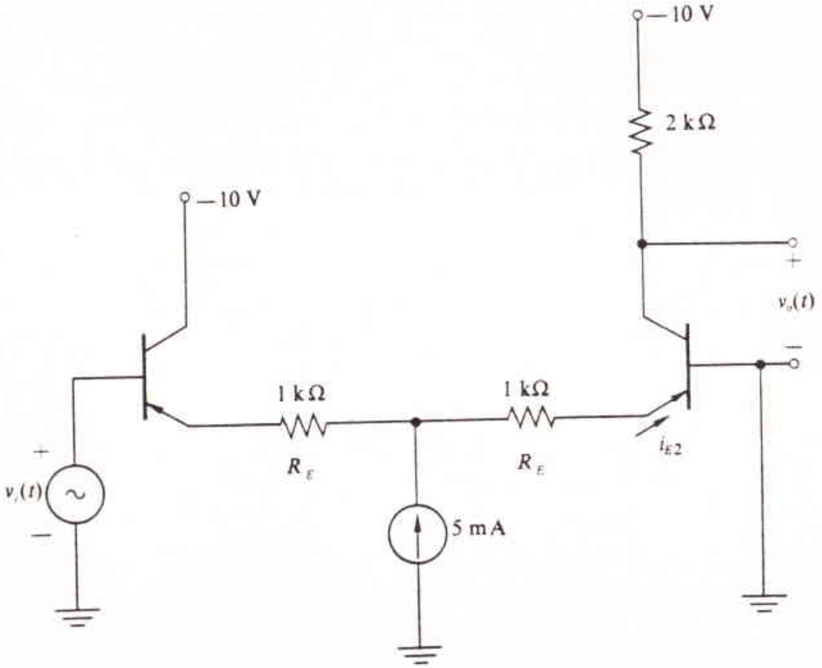


Figure 4.P–12

APPENDIX TO CHAPTER 4

FOURIER EXPANSIONS

The general approach to finding the average value or any harmonic component of a repetitive pulse train of period $T = 2\pi/\omega$ is to make a Fourier series expansion of it. In making such expansions, we choose the time origin so as to minimize terms and make use of various trigonometric identities to consolidate terms. The final form of the series is quite dependent on the variables chosen to express the waveshape and on the particular identities used to simplify the expression.

Most of the periodic waveforms $f(t)$ that we are interested in result from passing a dc-plus-cosine wave through a nonlinear nonmemory (no energy storage) device; hence the output waveforms possess symmetry about the $t = 0$ axis and thus may be expanded in a Fourier cosine series of the form

$$f(t) = a_0 + a_1 \cos \omega t + a_2 \cos 2\omega t + \cdots = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t,$$

where

$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} f(t)\,dt \tag{4.A–1}$$

and

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} f(t)\cos n\omega t\,dt. \tag{4.A–2}$$

It is usually convenient to change the independent variable from time to an angle in radians by defining $\theta = \omega t$ and then, since the waveshapes are symmetric about $\theta = 0$, to multiply by two and integrate from 0 to π. As functions of θ, a_0 and a_n take the forms

$$a_0 = \frac{1}{\pi}\int_0^{\pi} f\left(\frac{\theta}{\omega}\right) d\theta \tag{4.A–3}$$

and

$$a_n = \frac{2}{\pi}\int_0^{\pi} f\left(\frac{\theta}{\omega}\right)\cos n\theta\,d\theta. \tag{4.A–4}$$

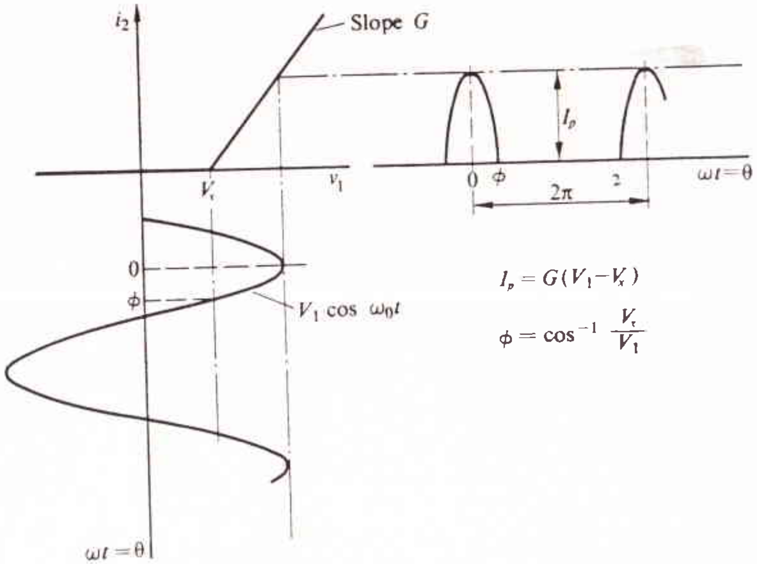


Fig. 4.A–1 Train of sine-wave tip current pulses.

Linear Sine-Wave Tips

For the train of sine-wave tip current pulses shown in Fig. 4.A–1 we observe that $i_2 = 0$ for $|\theta| \geq |\phi|$ (in the interval $\pi \leq \theta \leq \pi$), where $\phi = \cos^{-1}(V_x/V_1)$; hence the fundamental component I_1 of $i_2(t)$ may be written as

$$I_1 = \frac{2}{\pi}\int_0^{\phi} i_2\left(\frac{\theta}{\omega}\right)\cos\theta\,d\theta. \tag{4.A–5}$$

However, for $0 < \theta < \phi$, $i_2(t)$ has the value

$$i_2(t) = G(V_1 \cos \omega t - V_x);$$

therefore,

$$\begin{aligned} I_1 &= \frac{2}{\pi}\int_0^{\phi} G(V_1\cos\theta - V_x)\cos\theta\,d\theta \\ &= \frac{2G}{\pi}\left(\frac{V_1\phi}{2} + \frac{V_1 \sin 2\phi}{4} - V_x \sin\phi\right). \end{aligned} \tag{4.A–6}$$

By noting that $GV_1 = I_p/(1 - \cos\phi)$ and $GV_x = I_p\cos\phi/(1 - \cos\phi)$ and employing the identity $\sin 2\phi = 2\sin\phi\cos\phi$, we can further simplify the expression for I_1 to obtain

$$I_1 = \frac{I_p}{\pi}\,\frac{\phi - \cos\phi\sin\phi}{1 - \cos\phi}. \tag{4.A–7}$$

In a similar fashion, we obtain

$$I_0 = \frac{I_p}{\pi}\,\frac{\sin\phi - \phi\cos\phi}{1 - \cos\phi} \qquad (4.A\text{–}8)$$

and

$$I_n = \frac{2I_p}{\pi}\,\frac{\cos\phi\sin n\phi - n\sin\phi\cos n\phi}{n(n^2-1)(1-\cos\phi)}, \qquad n \geq 2. \qquad (4.A\text{–}9)$$

Normalized plots of I_n/I_p vs. conduction angle (2ϕ) are given in Fig. 4.2–3 for values of n from zero to five. Negative values of the coefficients in the plot are indicated by a minus sign after the designation of the particular curve; that is $n = 4^-$ implies that the fourth harmonic is negative.

For small values of ϕ we can approximate

$$\sin\phi \approx \phi - \frac{\phi^3}{3!} + \frac{\phi^5}{5!} \qquad \text{and} \qquad \cos\phi \approx 1 - \frac{\phi^2}{2!} + \frac{\phi^4}{4!}$$

to obtain

$$\frac{I_0}{I_p} \approx \frac{1}{\pi}\left(\frac{2\phi}{3} - \frac{\phi^3}{90}\right), \qquad \frac{I_1}{I_p} \approx \frac{1}{\pi}\left(\frac{4\phi}{3} - \frac{7\phi^3}{45}\right),$$

and

$$\frac{I_n}{I_p} \approx \frac{1}{\pi}\,\frac{4\phi}{3}, \qquad n \geq 1. \qquad (4.A\text{–}10)$$

From Eq. (4.A–10) it is apparent that the ratio of I_n to I_0 approaches 2 as ϕ approaches zero. This, of course, is a property possessed by any narrow pulse train.

For a pulse as wide as 90°, however, $I_1/I_0 = 1.88$ and $I_2/I_0 = 1.56$; even for a 120° pulse, I_1/I_0 is 1.80, which is within 10% of the very narrow pulse ratio value of 2.

Square-Law Sine-Wave Tips

For the train of square-law sine-wave tip current pulses shown in Fig. 4.A–2, we observe that $i_2 = 0$ for $|\theta| \geq |\phi|$ (in the interval $-\pi \leq \theta < \pi$), where, as in the previous case, $\phi = \cos^{-1}(V_x/V_1)$. In addition, for $|\theta| < |\phi|$, $i_2(t)$ is given by

$$i_2(t) = \frac{I_{DSS}}{V_p^2}(V_1\cos\omega t - V_x)^2, \qquad (4.A\text{–}11)$$

or equivalently,

$$i_2(t) = I_p\left(\frac{V_1\cos\omega t - V_x}{V_1 - V_x}\right)^2 = I_p\,\frac{\cos\omega t - \cos\phi}{1 - \cos\phi}. \qquad (4.A\text{–}12)$$

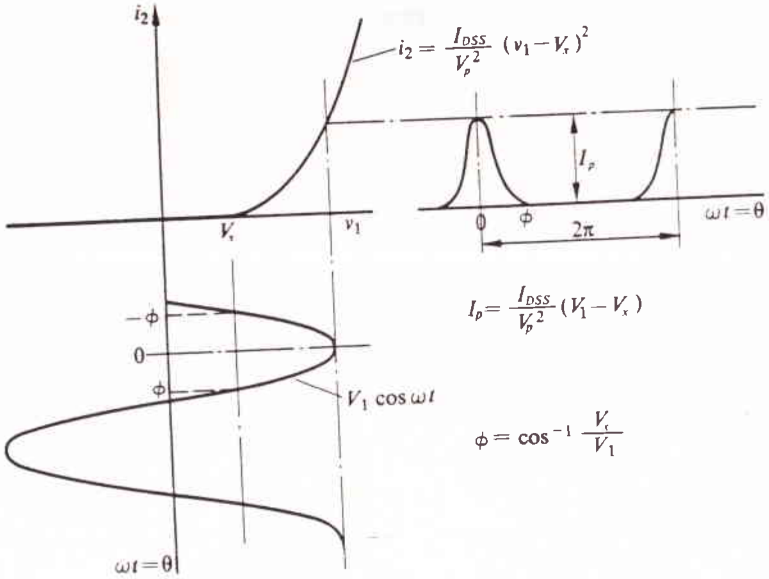

i_2
$i_2 = \frac{I_{DSS}}{V_p^2}(v_1 - V_x)^2$
V_x
v_1
I_p
0
ϕ
2π
$\omega t = \theta$
$I_p = \frac{I_{DSS}}{V_p^2}(V_1 - V_x)$
$-\phi$
0
ϕ
$V_1 \cos\omega t$
$\phi = \cos^{-1}\frac{V_x}{V_1}$
$\omega t = \theta$

Fig. 4.A–2 Train of square-law sine-wave tip current pulses.

where

$$I_p = \left(\frac{I_{DSS}}{V_p^2}\right)(V_1 - V_x)^2$$

is the peak value of i_2 and $\cos\phi = V_x/V_1$. Consequently, the fundamental component I_1 of $i_2(t)$ is given by

$$\begin{aligned} I_1 &= \frac{2I_p}{\pi}\int_0^{\phi}\frac{(\cos\theta - \cos\phi)^2\cos\theta\,d\theta}{(1-\cos\phi)^2} \\ &= \frac{2I_p}{\pi}\,\frac{\frac{3}{4}\sin\phi + \frac{1}{12}\sin 3\phi - \phi\cos\phi}{(1-\cos\phi)^2}. \end{aligned} \qquad (4.A\text{–}13)$$

In a similar fashion, we obtain

$$I_0 = \frac{I_p}{\pi}\,\frac{\phi - \frac{3}{4}\sin 2\phi + \frac{\phi}{2}\cos 2\phi}{(1-\cos\phi)^2}.$$

$$I_2 = \frac{2I_p}{\pi}\,\frac{\frac{\phi}{4} - \frac{1}{6}\sin 2\phi + \frac{1}{48}\sin 4\phi}{(1-\cos\phi)^2}. \qquad (4.A\text{–}14)$$

and

$$I_n = \frac{2I_p}{\pi}\,\frac{(4-n^2)\sin n\phi + (n-1)(n-2)\sin n\phi\cos 2\phi + 3n\sin(n-2)\phi}{(n^2-1)(n^2-4)n(1-\cos\phi)^2}, \qquad n > 2.$$

Normalized plots of I_n/I_p vs. conduction angle (2ϕ) are given in Fig. 4.4–4 for $n = 0, 1,$ and 2.

For small values of ϕ, we obtain the limiting forms of the Fourier coefficients given by

$$I_0 \approx \frac{I_p}{\pi}\frac{8\phi}{15} \qquad \text{and} \qquad I_n \approx \frac{2I_p}{\pi}\frac{8\phi}{15}. \tag{4.A–15}$$

Here, as in the case of the linear sine-wave tip, the sine and cosine terms are approximated by the first several terms in their MacLauren expansions. Again we observe that, for small values of ϕ (i.e., a small conduction angle), the ratio of I_n to I_0 approaches 2. In addition, we observe that, when the conduction angle (2ϕ) is as wide as 150°, $I_1/I_0 = 1.83$; hence for all narrower pulses $I_1/I_0 \approx 2$ is a good approximation.

CHAPTER 5

REACTIVE ELEMENT AND NONLINEAR ELEMENT COMBINATIONS

In this chapter we study the effect of combining nonlinear components with single reactances and with tuned circuits. We begin with the problem of capacitively coupling a periodic voltage to a general nonlinear load, and show that such coupling normally results in a dc bias shift which is a function of the ac input signal amplitude. Such an interrelationship rules out the use of superposition in such circuits. We consider both steady-state and transient operations of such circuits.

We then apply the solution of this general problem to the capacitively coupled transistor amplifier, where again we investigate both steady-state operation and transient operations. In addition, we obtain results not only for the transistor amplifier with a constant current bias but also for the resistively biased transistor amplifier. In both cases we obtain universal curves which may be applied to any transistor.

We then study the effect of nonlinear loading on a tuned circuit. We show that, if the circuit Q is sufficiently high, then the nonlinear load functions as an equivalent linear load which can be readily determined graphically.

Finally we consider the case where low-index AM signals are applied to a "clamping"-type circuit, and write a transfer function which is useful in later chapters.

For the reader who has not previously been exposed to this or similar material we recommend that he not try to absorb this chapter completely before proceeding to the subsequent chapters. Specifically, if he restricts himself to the steady-state phenomena of this chapter as well as the material in Section 5.5, he will be prepared to understand most of the material of the remaining chapters. He should then return to the transient portion of this chapter when the need arises.

5.1 CAPACITIVE COUPLING TO NONLINEAR LOAD

In many applications we must capacitively couple an ac driving source to a nonlinear load as shown in Fig. 5.1–1. Such coupling is usually required to block any dc component of the driving source from the load. This situation arises if $v_i(t)$ represents the output of one stage of a large-signal amplifier while the nonlinear load represents the input to the subsequent stage.

The basic difficulty in analyzing the apparently simple circuit of Fig. 5.1–1 is that, because of the nonlinear load, superposition is no longer valid. Thus the time-varying and dc components of the currents and voltages within the circuit are in general related. This interrelationship makes a general closed-form solution for

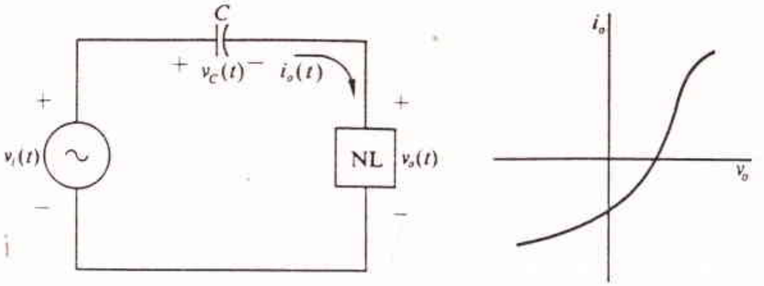


Fig. 5.1–1 Capacitive coupling to a nonlinear load.

$v_o(t)$, $v_C(t)$, and $i_o(t)$ in terms of $v_i(t)$ practically impossible. However, if we assume that $v_i(t)$ is periodic and that the series coupling capacitor C is an effective short circuit compared with the nonlinear load over the frequency range occupied by $v_i(t)$ [that is, that $v_C(t)$ contains only a dc component V_C, or at worst a slowly varying component compared with $v_i(t)$], then in principle we may determine the steady-state expressions for $v_o(t)$, $v_C(t)$, and $i_o(t)$. In addition, in most cases, we are able to calculate the transient build-up to steady state if $v_i(t)$ is applied at $t = 0$. Fortunately, in most cases of practical interest, the above assumptions are valid.

In the time domain the assumption that the capacitor is an effective short circuit corresponds to the assumption that the circuit time constants are such that the capacitor voltage remains essentially constant over a cycle of $v_i(t)$. Consequently, if $v_i(t)$ is applied at $t = 0$, the capacitor voltage slowly adjusts itself to its steady-state value $v_C = V_C = -V_{dc}$, and in turn $v_o(t)$ approaches $v_i(t) + V_{dc}$. Since in the steady state the average value of the capacitor current must equal zero, and since the capacitor current and nonlinear load current are equal, $V_C = -V_{dc}$ may be found as that value which causes the average value of the load current to equal zero.

Piecewise-Linear v_o-i_o Characteristic

If the v_o-i_o characteristic is piecewise-linear in form, we can determine $V_C = -V_{dc}$ by the following method. We assume a value of V_C, apply $v_o(t) = v_i(t) - V_C$ to the nonlinear characteristic, and with the techniques of Chapter 4 determine $i_o(t)$ and the average value of $i_o(t)$, that is, $\bar{i}_o = I_C$. We then repeat this procedure until we obtain enough values to plot a smooth curve of V_C vs. I_C. The point at which the curve intersects the $I_C = 0$ axis determines the steady-state value of $V_C = -V_{dc}$, from which the steady-state values of $v_o(t)$ and $i_o(t)$ follow directly.

As an example of this method, consider the nonlinear characteristic to be described by the piecewise-linear curve of Fig. 5.1–2 and $v_i(t) = (3\text{ V})\cos \omega t$. Figure 5.1–2 also indicates two additional characteristics which combine in parallel to yield the original characteristic. [This superposition of characteristics is performed only to focus attention on the structure of $i_o(t)$.]

If we now assume that the capacitor C is an ac short at ω and that its voltage is given by V_C, then

$$v_o(t) = 3\text{ V}\cos \omega t - V_C$$

and

$$i_o(t) = i_{o1}(t) + i_{o2}(t),$$

where

$$i_{o1}(t) = \frac{v_o(t) - (2\text{ V})}{2\text{ k}\Omega}$$

and $i_{o2}(t)$ is a periodic train of sine-wave tip pulses of peak amplitude

$$I_P = \frac{(3\text{ V}) - (2\text{ V}) - V_C}{2\text{ k}\Omega}$$

and conduction angle

$$2\phi = 2\cos^{-1}\frac{V_C + (2\text{ V})}{3\ V}.$$

The average capacitor current I_C is given by

$$I_C = \bar{i}_o = \bar{i}_{o1} + \bar{i}_{o2},$$

where,

$$\bar{i}_{o1} = -\frac{-V_C - 2\text{ V}}{2\text{ k}\Omega}$$

and $\bar{i}_{o2}$ is the average value of the train of sine-wave tips, which can be determined directly from Fig. 4.2–4 once a value of V_C is assumed ($V_x = V_C + 2\text{ V}$). Table 5.1–1 tabulates values for $\bar{i}_{o1}$, $\bar{i}_{o2}$, and I_C for several values of V_C. These data are also plotted in Fig. 5.1–3, from which it can be observed that $I_C = 0$ for $V_C = -1.33$ V. Consequently, in the steady state

$$V_C = -1.33\text{ V} \qquad \text{and} \qquad v_o(t) = -(3\text{ V})\cos \omega t + (1.33\text{ V}).$$

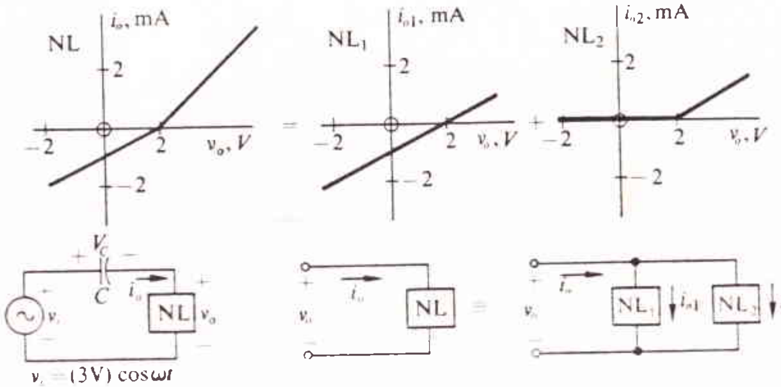


Fig. 5.1–2 Nonlinear characteristic for the circuit of Fig. 5.1–1.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION Y BIBLIOTECA
MONTEVIDEO - URUGUAY

Table 5.1–1 Tabulation of $\bar{i}_{o1}$, $\bar{i}_{o2}$, and I_C vs. V_C

V_C, V	$\bar{i}_{o1}$, mA	$\bar{i}_{o2}$, mA	$I_C = \bar{i}_{o1} + \bar{i}_{o2}$, mA
−5	1.5	1.5	3
−4	1	1.09	2.09
−3	0.5	0.76	1.26
−2	0	0.48	0.48
−1	−0.5	0.26	−0.24
0	−1	0.09	−0.91
1	−1.5	0	−1.5
2	−2	0	−2

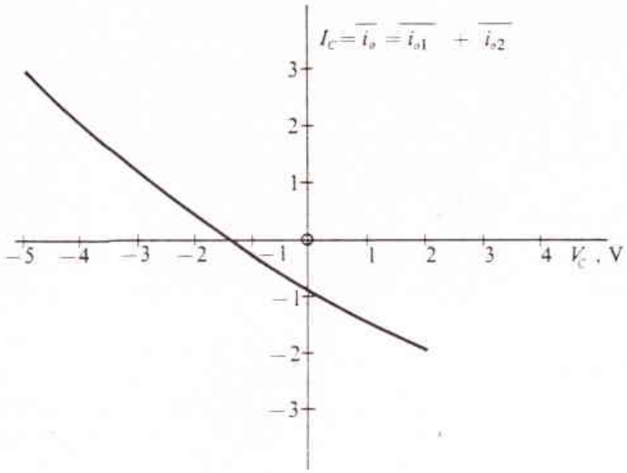


Fig. 5.1–3 Plot of I_C vs. V_C.

Actually, one should see by inspection that only two points need to be computed to solve such a problem. Since the right-hand branch in Fig. 5.1–2 contains an ideal diode, $\bar{i}_{o2}$ must be positive; hence for $\bar{i}_{o1}$ (in a branch with only a resistance and a battery) to just balance it, the current in the left-hand branch must be negative. Therefore, V_C must be greater than −2 V to make $\bar{i}_{o1}$ negative. On the other hand, if V_C exceeds zero, then $\bar{i}_{o2}$ will be zero, so that two reasonable initial choices for V_C are −2 V and −1 V. Since I_0 is a relatively slowly varying function of voltage in the region of interest, straight-line approximations for both currents will be reasonable; hence, a simple interpolation is possible to obtain $V_C \approx -1.33$ V.

This problem focuses attention on the clamping phenomenon which occurs in capacitively coupled nonlinear circuits. The term clamping refers to the shift in the average value of the load voltage v_o with drive level. With $v_i(t) = 0$, in the steady state $\bar{v}_o = 2$ V; hence, as the amplitude of v_i increases from 0 to 3 V, the average value

of $v_o(t)$ shifts from 2 V to 1.33 V. A further increase in the level of $v_i(t)$ causes the average value of $v_o(t)$ to decrease still further. We thus see dramatically that in such a nonlinear circuit the total output voltage *cannot* be obtained by the superposition of independent ac and dc components, as it can be in a linear circuit.

Analytic v_o-i_o Characteristic

If the v_o-i_o characteristic is described by an analytic relationship, it is possible that V_C, $v_o(t)$, and $i_o(t)$ may be obtained without resorting to a graphical analysis. For example, consider the circuit shown in Fig. 5.1–4, for which

$$i_o(t) = I_S e^{qv_o/kT} - I_{dc}. \tag{5.1–1}$$

For this case, if we assume that $v_i(t) = V_1 \cos \omega t$ and that C is a short circuit at the frequency ω such that $v_C(t) = V_C = -V_{dc}$, then

$$i_o(t) = I_S e^{qV_C/kT} e^{x \cos \omega t} - I_{dc}, \tag{5.1–2}$$

where, as usual, $x = qV_1/kT$. In addition,

$$\bar{i}_o = I_C = I_S e^{qV_C/kT} I_0(x) - I_{dc}, \tag{5.1–3}$$

where $I_0(x)$ is the modified Bessel function of order zero. Hence, in steady state, where $I_C = 0$, we have

$$\begin{aligned} V_{dc} = -V_C &= \frac{kT}{q} \ln \frac{I_{dc}}{I_S} - \frac{kT}{q} \ln I_0(x) \\ &= V_{dcQ} - \frac{kT}{q} \ln I_0(x) = V_{dcQ} - \Delta V, \end{aligned} \tag{5.1–4}$$

where V_{dcQ} is the value of V_{dc} with V_1 (or x) equal to zero and ΔV is the bias depression (or the variation in clamping voltage) induced by the driving voltage. Thus, for the given non-linearity and drive function, V_{dc} is a simple analytic function of I_{dc} and x (or V_1).

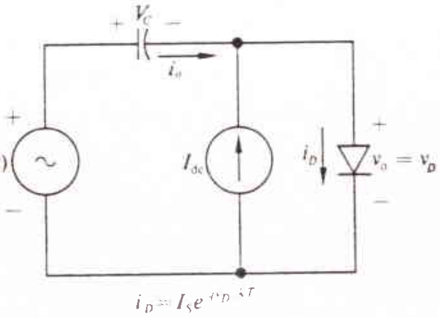


Fig. 5.1–4 Nonlinear block consisting of a current generator and a diode.

For values of x larger than 4($V_1 \geq 104$ mV), one may approximate the $I_0(x)$ Bessel function term to within better than 4% (within 6% for the $x \geq 3$) by

$$I_0(x) \approx \frac{e^x}{\sqrt{2\pi x}}. \qquad (5.1\text{–}5)$$

from which it follows that

$$\frac{kT}{q} \ln I_0(x) = V_1 - \frac{kT}{2q} \ln 2\pi x = V_1 - (23.8 \text{ mV}) - (13 \text{ mV}) \ln x.$$

Therefore, for $x > 4$,

$$V_{dc} = V_{dcQ} - V_1 + (23.8 \text{ mV}) + (13 \text{ mV}) \ln x. \qquad (5.1\text{–}6)$$

Since the (13 mV) ln x term varies by less than 32 mV as x varies from 5 to 50 (130 to 1300 mV), one sees that V_{dc} will vary almost in a straight-line fashion with V_1 over this region. The value of V_{dcQ} will be a function of both I_S (a junction parameter) and I_{dc} (a circuit parameter). Since the natural logarithm of 10 is 2.3, a 10 to 1 change in one of these quantities will result in only a 60 mV change in V_{dcQ}. Figure 5.1–5 plots V_{dc} vs. x and V_1 for the case where $I_S = e^{-30}$ mA $= 10^{-16}$ A (a typical value for an integrated-circuit silicon diode) and for I_{dc} values of 0.1, 1.0, and 10.0 mA.

For $I_{dc} = 1.0$ mA,

$$V_{dcQ} = 30 \times 26 = 780 \text{ mV}.$$

Hence for $x > 4$, one can approximate V_{dc} quite closely as

$$(780 + 24 + 26) \text{ mV} - V_1 \qquad \text{or} \qquad (830 \text{ mV}) - V_1.$$

What this implies is that as far as calculating voltage levels is concerned, the constant current source and actual exponential diode combination may be replaced by an ideal diode in series with a battery as shown in Fig. 5.1–6.

This ideal clamping circuit clamps the top of V_1 to the positive battery voltage; hence $V_{dc} = -V_C$ is always just the battery voltage minus V_1. Therefore, to within the precision allowed by the approximation, peak values of driving voltage greater than 830 mV will result in negative values of V_{dc}, while small values will result in positive values of V_{dc}. Once $V_{dc} = V_C$ is known, it may be inserted into Eq. (5.1–2) to determine $i_o(t)$.

As an example of using the model of Fig. 5.1–6, consider the circuit shown in Fig. 5.1–7. For this circuit, we first replace the LR combination by a current-source approximation and then use the model of Fig. 5.1–6 to arrive, in a simple fashion, at the output ac signal voltages and dc output level. Once the output voltages are known, one can revert to the basic diode equation or to the curves of Chapter 4 to determine the diode or source currents.

Consider the practical case where the reactance of L is large with respect both to the loss term and to the equivalent diode impedance, so that the ac current through the inductive branch may be assumed to be negligible.

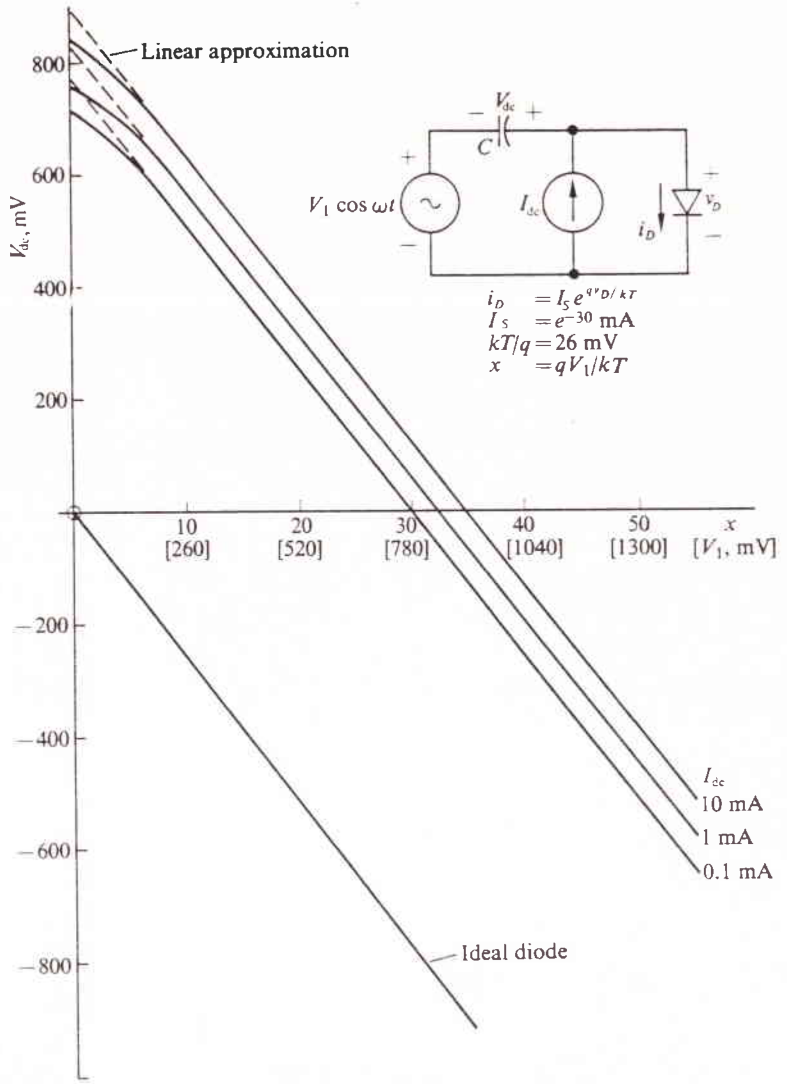


Fig. 5.1–5 Plot of V_{dc} vs. x and V_1.

Then from the viewpoint of dc current, the inductor-resistor combination may be replaced by a constant-current generator. Then the circuit of Fig. 5.1–7 may be replaced by the circuit of Fig. 5.1–4 and finally by the model of Fig. 5.1–6. In this circuit, the final value of I_{dc} is also constrained by the additional dc circuit relationship

$$I_{dc} r = V_C.$$

If we assume that $I_S = e^{-30}$ mA (as in Fig. 5.1–5), then with a drive voltage $V_1 = 425$ mV an assumption of $I_{dc} = 1$ mA would lead to a battery voltage in the

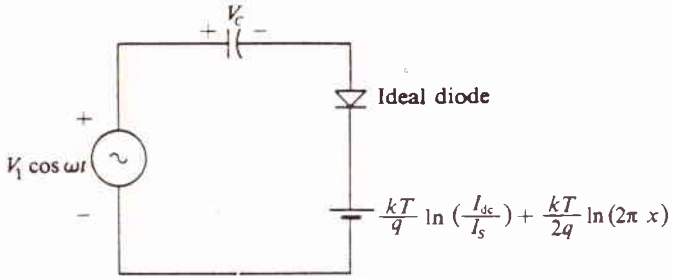

V_C
+
−
Ideal diode
+
$V_1 \cos \omega t$
−
$\frac{kT}{q} \ln\left(\frac{I_{dc}}{I_s}\right) + \frac{kT}{2q} \ln(2\pi\, x)$

Fig. 5.1–6 Voltage-clamping model for exponential diode and current source combination.

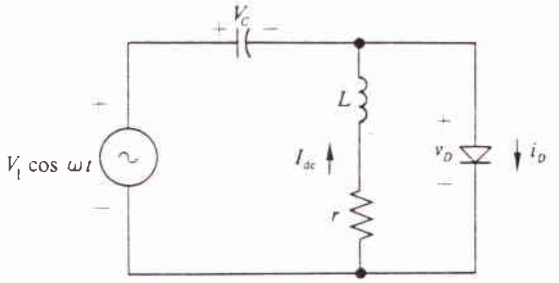


Fig. 5.1–7 Inductively biased clamping ciruit.

model of Fig. 5.1–6 of 840 mV. (This is 10 mV more than the value used previously, because we used the actual value of x instead of an estimated value.) Since $V_C = 840 - 425 = 415$ mV, then if $r = 500\ \Omega$ (a possible value for a 1 H torroidal coil) the circuit relationship would lead to $I_{dc} = 415/500 = 0.83$ mA.

Now the model could be recalculated for the $I_{dc} = 0.83$ mA case instead of for the 1 mA assumed initially. If this is done, the change in the battery voltage will be less than 5 mV out of 840 mV or much less than 1 %; hence such an additional cycle of calculation is probably not warranted.

If $r = 50\ \Omega$, then our initial calculation would lead to $I_{dc} = 8.3$ mA. A recalculation would lead to an increase in the battery voltage of $26 \ln 8.3 = 55$ mV and hence in I_{dc} of 1.1 mA = (55/50) mA. A further recalculation would only increase the battery voltage by $26 \ln (9.4/8.3) = 3.2$ mV or less than 0.4 % and hence would rarely seem necessary.

In the real world one seldom has the simple case shown in Fig. 5.1–4. In practical circuits one often has a driving voltage source with a nonnegligible source impedance while at the same time the diode (or transistor junction) may have either a deliberately added or an inherent series resistive component.

The circuit that we shall consider is shown in Fig. 5.1–8. (The case of a finite source impedance and a resistanceless diode may be transformed into the form of Fig. 5.1–8 by adding equivalent current sources to "move" the I_{dc} generator to the other end of the source resistance. The amount of the variation in the dc voltage across the capacitor is identical in the two cases.)

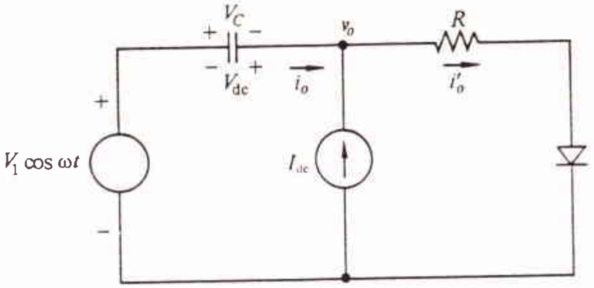


Figure 5.1–8

In this case

$$v_o(t) = i'_o(t)R + \frac{kT}{q} \ln\left[\frac{i'_o(t)}{I_{dc}}\right] + \frac{kT}{q} \ln\left(\frac{I_{dc}}{I_{ES}}\right) \tag{5.1–7}$$

$$= V_1 \cos \omega t - V_C .$$

Now the capacitor will not pass dc, hence $\overline{i_o(t)} = 0$ or equivalently $\overline{i'_o(t)} = I_{dc}$. The circuit will adjust itself to the value of dc capacitor voltage that causes $\overline{i'_o(t)} = I_{dc}$.

A general relationship between V_C and $\overline{i_o(t)}$ does not seem possible; however, solutions are possible for the two limiting cases where $I_{dc}R \ll kT/q$ and where $I_{dc}R \gg kT/q$. (The second case is equivalent to $g_{in}\ R \gg 1$.)

On the basis of our experience with the curves of Section 4.8, we plot the bias depression ΔV for these two extreme cases, normalized as in Section 4.8, and find that they form a tight set. We then conjecture that the $I_{dc}R = kT/q$ case falls halfway between the other two; thus we have the wherewithal to estimate $V_C = V_{CQ} + \Delta V$ for all practical cases.

When R is zero we revert to Eq. (5.1–4) to form $\Delta V = V_C - V_{CQ} = (kT/q) \ln I_0(x)$ and normalize it to V_{co}, which is kT/q in the $R = 0$ case.

When $I_{dc}R \gg kT/q$, the physical diode resistor combination may be replaced by the model shown in Fig. 4.8–3 to obtain the circuit shown in Fig. 5.1–9. For this circuit the quiescent value of $-V_{CQ} = V_{dcQ}$ is $I_{dc}R + V_0$, and the normalizing voltage

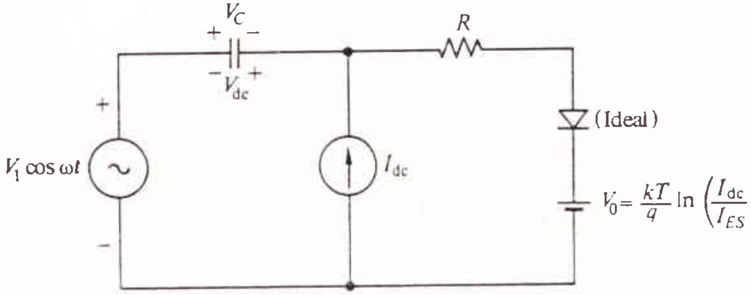


Figure 5.1–9

$V_{co} = I_{dc}R$. In addition the ideal diode remains forward biased for $V_1/V_{co} \leq 1$, the circuit operates linearly, and no bias depression $\Delta V = V_{dcQ} - V_{dc}$ exists.

For $V_1/V_{co} > 1$, i_o becomes a periodic train of sine-wave tip pulses and V_{dc} adjusts itself to cause $\bar{i_o} = I_{dc}$. This is precisely the problem that was solved in Section 4.8, from which we obtain that

$$V_C = V_1 \cos\phi - V_0 = -V_{dc} \tag{5.1-9}$$

and

$$\frac{V_{dcQ} - V_{dc}}{V_{co}} = \frac{\Delta V}{V_{co}} = 1 + \frac{V_1}{V_{co}}\cos\phi. \tag{5.1-10}$$

Thus V_C and in turn $\Delta V/V_{co}$ can be evaluated in terms of V_1 and V_{co} from the values of ϕ given in Table 4.8-1.

Now both these curves, normalized with respect to

$$V_{co} = \frac{kT}{q} + I_{dc}R = I_{dc}(r_{in} + R),$$

are shown in Fig. 5.1-10. The dotted curve halfway between the extreme cases is the conjectured curve for $I_{dc}R = kT/q$ or $R = r_{in}$.

As an example of the above analysis, let $I_{dc} = 2.6$ mA, $R = 10$ ohms, $V_1 = 520$ mV, and $V_0 = kT/q \ln(I_{dc}/I_{ES}) = 0.65$ V in the circuit of Fig. 5.1-8. With these values we observe that $V_{CQ} = -0.676$ V, $g_{in}R = 1$, $V_{co} = I_{dc}(R + r_{in}) = 52$ mV and $V_1/V_{co} = 10$; hence from Fig. 5.1-10 we obtain $\Delta V/V_{co} \approx 7$, from which follows

$$V_C = (-0.676 \text{ V}) + (7)(52 \text{ mV}) = -0.312 \text{ V}$$

and

$$v_o(t) = (312 \text{ mV}) + (520 \text{ mV}) \cos\omega t.$$

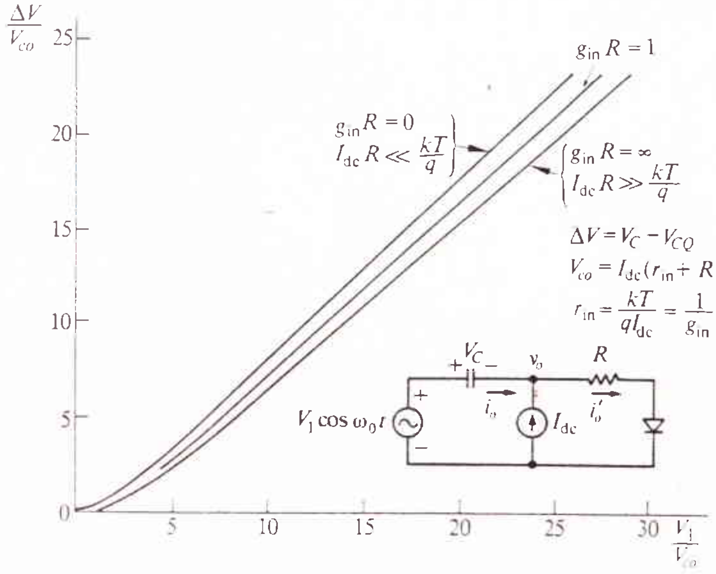


Figure 5.1-10

With this value of $v_o(t)$ one can return to the characteristics of Fig. 4.8-2 and obtain a reasonably accurate point-by-point plot of the diode-resistor current. For example $v_o(t) - V_0$ has a peak value of 282 mV; thus $(v_o(t) - V_0)_{max}/V_{co} \approx 5.4$ in this example. This value determines a normalized peak current from Fig. 4.8-2 of $I_{op}/I_{dc} \approx 8.6$ or $I_{op} \approx 22.4$ mA.

With $R = 0$ but with the same drive voltage, and the same I_{dc} and I_{ES}, the capacitor dc voltage would be +193 mV, while the peak current would be approximately 29 mA or 30% larger. Hence an apparently small source resistance or series resistor can have a reasonably large effect on both dc, ac, and peak values of the circuit's voltages and currents. The key to the effect that the resistor will have is the size of $I_{dc}R$ with respect to $(kT/q) = 26$ mV or equivalently the size of $g_{in}R$ with respect to 1.

5.2 TRANSIENT BUILD-UP TO STEADY STATE

If $v_i(t)$ in Fig. 5.1-1 is periodic and applied to the circuit at $t = 0$, we can calculate the transient build-up of V_C and $v_o(t)$ to their steady-state values computed above. Clearly in this case $V_C = V_C(t)$ is a function of time; however, in order to obtain a reasonable solution, we continue to assume that this function of time is varying sufficiently slowly so that $V_C(t)$ is essentially constant over any complete cycle of $v_i(t)$. With this assumption, $v_o(t)$, $i_o(t)$, and $I_C = I_C(t)$ can be evaluated directly in terms of $V_C(t)$ by the same method employed in obtaining their steady-state values. In particular, the V_C–I_C curve obtained to determine the steady-state value of $V_C = V_{CSS}$ (cf. Fig. 5.1-3) now relates the slowly varying capacitor current $I_C(t)$ to the capacitor voltage $V_C(t)$ by an expression of the form

$$I_C(t) = f(V_C(t)). \tag{5.2-1}$$

The capacitor itself relates $I_C(t)$ to $V_C(t)$ by the expression

$$I_C(t) = C\frac{dV_C(t)}{dt}; \tag{5.2-2}$$

hence, by eliminating $I_C(t)$ from Eqs. (5.2-1) and (5.2-2), we obtain the first-order nonlinear differential equation for $V_C(t)$ in the form

$$C\frac{dV_C}{dt} = f(V_C). \tag{5.2-3}$$

If $f(V_C)$ is known in analytic form, we can solve Eq. (5.2-3) directly, as we illustrate in Section 5.3. On the other hand, we can expand $f(V_C)$ in a Taylor series about the eady-state value of $V_C = V_{CSS}$ to obtain

$$C\frac{dV_C}{dt} = f(V_{CSS}) + \frac{\partial f(V_{CSS})}{\partial V_C}(V_C - V_{CSS}) + \frac{\partial^2 f(V_{CSS})}{\partial V_C^2}\frac{(V_C - V_{CSS})^2}{2!} + \cdots$$
$$= -G(V_C - V_{CSS}) + b_2(V_C - V_{CSS})^2 + \cdots, \qquad (5.2\text{–}4)$$

here $I_{CSS} = f(V_{CSS}) = 0$, $G = -\partial f(V_{CSS})/\partial V_C$, and

$$b_n = \frac{1}{n!}\frac{\partial^n f(V_{CSS})}{\partial V_C^n} \qquad \text{for } n \geq 2.$$

he advantage of the Taylor series expansion is that, in general, only a few terms must e kept to closely approximate $f(V_C)$ over the range of interest of V_C; in addition, if (V_C) is available only graphically, as in Fig. 5.1–3, the coefficients of the first few erms of the Taylor series can be determined by standard curve-fitting techniques to losely approximate the V_C–I_C curve.

As a first approximation to the transient build-up, we keep only the first nonzero erm in the Taylor expansion to obtain

$$C\frac{dV_C}{dt} = -G(V_C - V_{CSS}), \qquad (5.2\text{–}5)$$

vhere $-G$ is the slope of the I_C–V_C curve evaluated at $V_C = V_{CSS}$. Equation (5.1–5) s linear, with the solution

$$V_C(t) = \{V_{CSS} + [V_C(0^+) - V_{CSS}]e^{-t/\tau}\}u(t), \qquad (5.2\text{–}6)$$

vhere $\tau = C/G$.

For the nonlinearity of Fig. 5.1–2, which results in the V_C–I_C curve shown in ig. 5.1–3, if $v_i(t) = (3\text{ V})\cos\omega t$ is applied at $t = 0$, then

$$V_C(0^+) = -2\text{ V}, \qquad V_{CSS} = -1.33\text{ V}, \qquad \text{and} \qquad G = 0.72\text{ m}\Omega;$$

ience $V_C(t) = [-(1.33\text{ V}) - (0.67\text{ V})e^{-t/\tau}]u(t)$ and $v_o(t) = [-V_C(t) + (3\text{ V})\cos\omega t]u(t)$, vhere $\tau = (1.39\text{ k}\Omega)C$. A sketch of $v_o(t)$ is shown in Fig. 5.2–1. Since $V_C(t)$ varies nly between -2 V and -1.33 V, the linear approximation for the V_C-I_C characeristic is quite valid.

When terms beyond the first term of the Taylor series expansion for $f(V_C)$ are kept, then the solution to Eq. (5.2–4) is obtained in the following manner. We rearrange the equation in the form

$$\frac{-Gdt}{C} = \frac{dV_C}{(V_C - V_{CSS}) - (b_2/G)(V_C - V_{CSS})^2 \cdots}, \qquad (5.2\text{–}7)$$

we expand the denominator of the right-hand side in partial fractions, and we integrate term by term. We illustrate this procedure for the case where $f(V_C)$ may be approximated by the first two nonzero terms in its Taylor expansion, i.e., where

$$C\frac{dV_C}{dt} = -G(V_C - V_{CSS}) + b_2(V_C - V_{CSS})^2. \qquad (5.2\text{–}8)$$

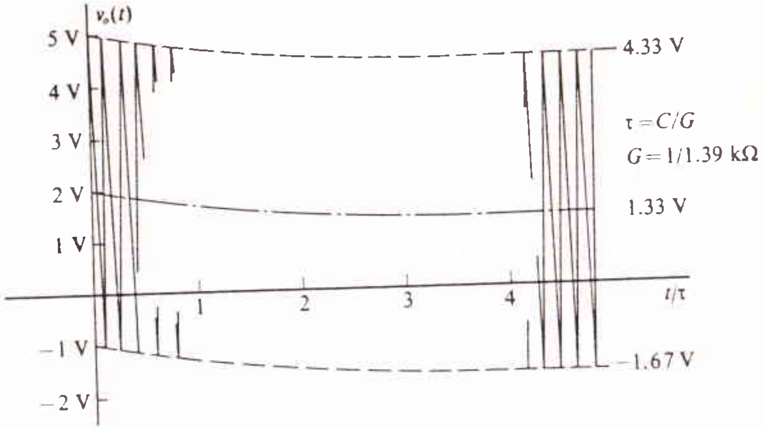


Fig. 5.2–1 Output transient when a 3 V peak sine wave is suddenly applied to the circuit of Fig. 5.1–2.

For this case,

$$-\frac{Gdt}{C} = \frac{dV_C}{(V_C - V_{CSS})[1 - (b_2/G)(V_C - V_{CSS})]}$$
$$= \frac{dV_C}{V_C - V_{CSS}} + \frac{(b_2/G)\,dV_C}{1 - (b_2/G)(V_C - V_{CSS})}. \qquad (5.2\text{–}9)$$

Upon integrating and rearranging terms, we obtain

$$V_C(t) = V_{CSS} + \frac{Ke^{-t/\tau}}{1 + (b_2/G)Ke^{-t/\tau}}, \qquad (5.2\text{–}10)$$

where K is a constant of integration and $\tau = C/G$. When b_2 is zero, this equation of course reduces to the linear case. Unless Kb_2/G is appreciable with respect to unity, the influence of the square term is always small. In any case, b_2 does not influence the length of the transient. In most cases b_2 affects merely the initial slope and the values of $V_C(t)$ for $t < \tau$.

A similar result can be obtained if $f(V_C)$ consists of a linear plus a cubic term. Specifically, if Eq. (5.2–4) has the form

$$C\frac{dV_C}{dt} = -G(V_C - V_{CSS}) + b_3(V_C - V_{CSS})^3, \qquad (5.2\text{–}11)$$

then the solution for $V_C(t)$ takes the form

$$V_C(t) = V_{CSS} + \frac{Ke^{-t/\tau}}{\sqrt{1 + (b_3/G)Ke^{-2t/\tau}}}, \qquad (5.2\text{–}12)$$

where K is a constant of integration and $\tau = C/G$. Again, unless b_3K/G is appreciable compared to unity, the b_3-term has no influence. And again the time of the transient is set entirely by $\tau = C/G$, as it is for any polynomial approximation, of whatever order, of $f(V_C)$ or for an exact analytic expression for $f(V_C)$.

For the diode nonlinearity of Fig. 5.5–1, $I_C = f(V_C)$ is given by Eq. (5.1–3) and

$$G = \frac{\partial f(V_{CSS})}{\partial V_C} = \frac{qI_S}{kT} e^{qV_{CSS}/kT} I_0(x) = \frac{qI_{\text{dc}}}{kT}. \tag{5.2–13}$$

As we shall see in the next section, the transient is governed by the time constant (cf. Eq. 5.3–11)

$$\tau = C/G = \frac{kTC}{qI_{\text{dc}}}.$$

With a knowledge of the rate at which V_C varies, we are now in a position to determine the value of C which justifies our assumption that $V_C(t)$ is constant over a cycle of $v_i(t)$. If the period of $v_i(t)$ is T, and if $\tau > 20T$, then $V(t)$ varies less than 5% in the time interval T. Consequently, if

$$C > 20GT = \frac{40\pi G}{\omega}$$

or

$$\frac{\omega C}{G} > 40\pi, \tag{5.2–14}$$

where $\omega = 2\pi/T$, then our assumptions are justified. In many practical situations the inequality of Eq. (5.2–14) may be relaxed by a factor of 2 or 3 without any noticeable effect on the expressions derived.

5.3 CAPACITIVELY COUPLED TRANSISTOR AMPLIFIER—CONSTANT CURRENT BIAS

Since a great many large-signal ac transistor amplifiers are capacitively coupled, we shall study this coupling problem in detail in this section. We begin by analyzing the capacitively coupled amplifier shown in Fig. 5.3–1, which is biased from a constant current source I_{dc} and driven at its emitter terminal (common base) by the input voltage $v_i(t)$. We then show, in Section 5.4, that the analysis is essentially unchanged if the drive is applied to the base with the emitter capacitively bypassed (common emitter) or if the bias is developed by the more conventional scheme of resistively dividing $-V_{CC}$ and applying it to the base with a resistor R_E placed in the emitter circuit. In all cases, we assume, however, that the input volt-ampere characteristic of the transistor may be modeled by an exponential relationship of the form

$$i_E = I_{ES} e^{qv_{EB}/kT} = I_{ES} e^{-qv_{BE}/kT}, \tag{5.3–1}$$

that the collector-to-emitter voltage never reaches zero and saturates the transistor, that the collector current may be modeled by a current source αi_E, where $\alpha \approx 1$, over the entire range of emitter currents, and that the internal reactances of the transistor, such as charge storage, may be neglected. These assumptions provide an excellent first-order approximation to the performance of many transistor amplifiers for values of $|Z_L|$ less than 100 kΩ and frequencies up to the tens of megahertz.

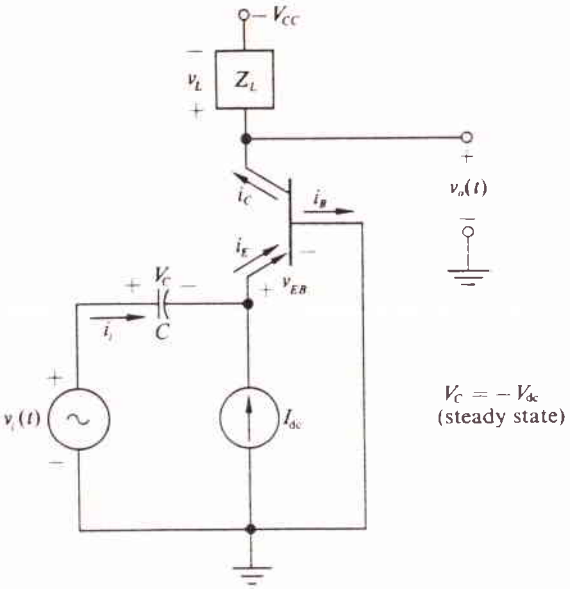


Fig. 5.3–1 Capacitively coupled transistor amplifier.

For the circuit of Fig. 5.3–1, we first determine its steady-state behavior with $v_i(t)$ a periodic drive and then determine how the circuit builds up to steady state. In addition, in Section 5.6 we develop the transfer function of the network for the case where $v_i(t)$ is an AM wave of low modulation index. This transfer function proves to be useful for evaluating the steady-state stability of transistor oscillators and the output signals of transistor limiters.

If we assume for the circuit of Fig. 5.3–1 that the coupling capacitor C is sufficiently large that its voltage v_C remains constant at $V_C = -V_{\text{dc}}$ over a cycle of $v_i(t)$, then the model of Fig. 5.1–6 is immediately applicable.

For a given input voltage, we determine the value of the battery and hence of the clamping level and the capacitor voltage. We then determine the emitter current (the diode current in the model) from the known dc value of current, V_C, and the driving voltage waveshape.

For example, assume that the waveshape of Fig. 5.3–2 is applied as $v_i(t)$ in the circuit of Fig. 5.3–1.

As a reasonable approximation to aid in fixing the battery voltage in the model, we might model the input as a 200 mV peak sine wave: hence $x \approx 7.7$. For such an x, the $(kT/2q)\ln 2\pi x$ term becomes approximately 50 mV. If we assume that I_{dc} is 1 mA and that I_S is 10^{-16} A, then

$$\frac{kT}{q}\ln\frac{I_{\text{dc}}}{I_{ES}} = V_{\text{dc}\,Q} = 760 \text{ mV}.$$
]

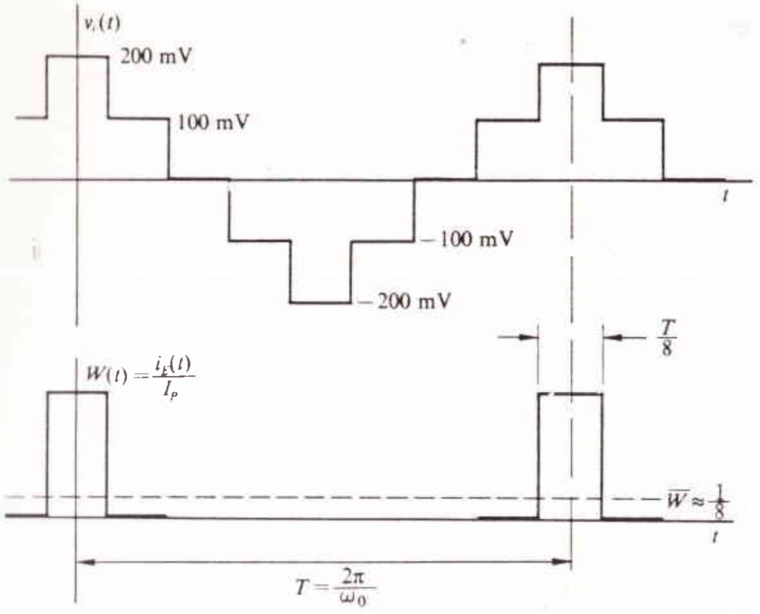

$v_i(t)$
200 mV
100 mV
t
−100 mV
−200 mV
$\frac{T}{8}$
$W(t) = \frac{i_E(t)}{I_P}$
$\overline{W} \approx \frac{1}{8}$
t
$T = \frac{2\pi}{\omega_0}$

Fig. 5.3–2 Plot of $v_i(t)$ and $i_E(t)/I_P = W(t)$ vs. t.

Ience, to a very good approximation, the emitter voltage will be clamped to a peak oltage of +810 mV; thus the capacitor voltage will be −610 mV. (If the 200 mV eak signal is to reach 810 mV, the capacitor voltage must add 610 mV to it.)

The emitter current can be expected to flow primarily during the most positive nterval of the driving signal.† Since this interval consumes only one-eighth of the otal period, the current during this interval must be eight times the average value, or mA.

Once the current is known to flow primarily in narrow pulses, it follows from hapter 4 that the fundamental component is approximately twice the dc value; hat is, $I_{E1} \approx 2I_{dc}$. Therefore, if Z_L is a high-Q parallel RLC circuit tuned at $= 2\pi/T$ with a parallel resistor R_L, then the output voltage $v_o(t)$ may be closely pproximated by

$$v_o(t) = -V_{CC} + 2\alpha I_{dc} R_L \cos \omega t. \tag{5.3–2}$$

If the stairstep waveshape were replaced by a 200 mV peak sine wave, then the ame clamping level and capacitor voltages would exist. For this case, the ratio of he peak current to the average current is

$$\frac{I_P}{I_{dc}} = \frac{e^x}{I_0(x)} \approx \sqrt{2\pi x} \tag{5.3–3}$$

---

† The currents that flow when the input to the emitter is 810 mV and when it is 710 mV are roportional to $e^{810/26}/e^{710/26} = e^{3.8} = 45$. Hence the current during the maximum voltage eriod is 45 times the current flow during the adjacent segments. Thus assuming that all the urrent flows in one flat-topped pulse will lead at most to about a 2% error.

for $x > 4$. Thus for $x = 7.7$ and $I_{dc} = 1$ mA, $I_P \approx 7$ mA. Figure 4.5–3 indicates the general shape of the current pulse, while Fig. 4.5–4 indicates a conduction angle of about 104°. Figure 4.5–5 indicates that for this value of x, $I_1 = 1.85 I_{dc}$; hence again, if the load is a tuned circuit, the output voltage can be found as

$$v_o(t) = -V_{CC} + 1.85\alpha I_{dc} R_L \cos \omega t. \tag{5.3–4}$$

If the load is tuned to some harmonic of the input driving waveshape, then either Fig. 4.5–5 or the tables of $2I_n(x)/I_0(x)$ that appear in the Appendix at the back of the book may be used to determine $v_o(t)$.

A more general approach to the circuit of Fig. 5.3–1 is to define a waveform function $W(t)$ as the ratio of the actual emitter current, $i_E(t)$, to the peak emitter current I_P. If the capacitor voltage has reached a steady-state value $V_C = -V_{dc}$, then

$$I_P = I_{ES} e^{qV_{dc}/kT} e^{qv_{i\max}/kT}, \tag{5.3–5}$$

$$i_E(t) = I_{ES} e^{qV_{dc}/kT} e^{qv_i(t)/kT}, \tag{5.3–6}$$

and

$$\begin{aligned} W(t) &= e^{qv_i(t)/kT} / e^{qv_{i\max}/kT} \\ &= i_E(t)/I_P. \end{aligned} \tag{5.3–7}$$

The waveform function contains the time-variable portion of $i_E(t)$; hence if I_P and $W(t)$ are known, then $i_E(t)$ is known. The average or dc value of $i_E(t)$ will be I_P times the average value of $W(t)$:

$$I_{dc} = \overline{i_E(t)} = I_P \overline{W(t)}. \tag{5.3–8}$$

Thus if I_{dc} and $v_i(t)$ are known, one first evaluates or plots $W(t)$, which is only a function of $v_i(t)$, and finds its average value $\overline{W(t)}$. If $\overline{W(t)}$ and I_{dc} are known, then I_P is fixed. However, fixing I_P fixes V_{dc} and completes the solution.

When an unbypassed series resistance exists in either the base or emitter circuits, the results of Sections 4.8 and 5.1 for the added resistance case are again directly applicable.

Transient Buildup

If the periodic input $v_i(t)$ is applied at $t = 0$ to the circuit of Fig. 5.3–1, during the buildup to steady state the slowly varying component of capacitor current is given by

$$\bar{i}_i = I_i(t) = I_{ES} e^{-qV_C(t)/kT} \overline{e^{qv_i(t)/kT}} - I_{dc} = C\frac{dV_C(t)}{dt}. \tag{5.3–9}$$

Hence the differential equation for $V_{dc}(t) = -V_C(t)$ takes the form

$$C\frac{dV_{dc}(t)}{dt} + I_{ES} e^{qV_{dc}(t)/kT} e^{qv_{i\max}/kT} \overline{W(t)} = I_{dc}. \tag{5.3–10}$$

We may integrate Eq. (5.3–10) in a straightforward fashion to obtain

$$e^{qV_{dc}(t)/kT} = \frac{(I_{dc}/I_{ES})e^{-qv_{i\,max}/kT}}{\overline{W(t)} + Ke^{-t/\tau}}, \tag{5.3–11}$$

where $\tau = CkT/qI_{dc} = C/g_{in}$, K is an arbitrary constant of integration, and $g_{in} = qI_{dc}/kT$ is the small-signal input conductance of the transistor in Fig. 5.3–1. If we assume that for t less than zero $v_i = 0$, then

$$e^{-qv_{i\,max}/kT} = \frac{I_{dc}}{I_{ES}}$$

and

$$K = e^{-qv_{i\,max}/kT} - \overline{W(t)}. \tag{5.3–12}$$

With this value of K, Eq. (5.3–11) reduces (for $t \geq 0$) to

$$e^{qV_{dc}(t)/kT} = \frac{(I_{dc}/I_{ES})e^{-qv_{i\,max}/kT}}{\overline{W(t)} + [e^{-qv_{i\,max}/kT} - \overline{W(t)}]e^{-t/\tau}}, \tag{5.3–13a}$$

or equivalently,

$$V_{dc}(t) = -v_{i_{max}} + \frac{kT}{q} \ln \frac{\dfrac{I_{dc}}{I_{ES}}}{\overline{W} + (e^{-qv_{i\,max}/kT} - \overline{W})e^{-t/\tau}}. \tag{5.3–13b}$$

Now, in the steady state, $\overline{W} = I_{dc}/I_P$ and $e^{-t/\tau} = 0$; thus Eq. (5.3–13) reduces to

$$V_{dc} = \underbrace{-v_{i_{max}} + \frac{kT}{q} \ln \frac{I_{dc}}{I_{ES}}}_{V_{dcQ}} - \frac{kT}{q} \ln \overline{W}. \tag{5.3–13c}$$

This exact form for V_{dc} also results if the waveform function is employed to evaluate the steady-state transistor response in a fashion discussed previously in this section.

By noting that

$$i_E(t) = I_{ES}e^{qV_{dc}(t)/kT}e^{qv_{i\,max}/kT}W(t)$$

and employing Eq. (5.3–13a), we may also obtain (for $t \geq 0$)

$$i_E(t) = \frac{I_{dc}[W(t)]}{\overline{W} + (e^{-qv_{i\,max}/kT} - \overline{W})e^{-t/\tau}}, \tag{5.3–14}$$

where, of course, $W(t) = e^{qv_i(t)/kT}/e^{qv_{i\,max}/kT}$. As $t \to \infty$, Eq. (5.3–14) reduces to the expected form $i_E = I_PW(t)$.

For $v_i(t)$ given by the staircase waveform of Fig. 5.3–2, $W(t)$ has the form shown in the same figure. $\overline{W} = \frac{1}{8}$, and $v_{i_{max}} = 200$ mV: hence, for this case, with $kT/q = 26$ mV and $I_{dc}/I_{ES} = 5 \times 10^{12}$,

$$i_E(t) = \frac{8I_{dc}[W(t)]}{1 + (3.7 \times 10^{-3} - 1)e^{-t/\tau}} \tag{5.3–15a}$$

and

$$V_{dc}(t) = (614\text{ mV}) - (26\text{ mV}) \ln [1 + (3.7 \times 10^{-3} - 1)e^{-t/\tau}]. \tag{5.3–15b}$$

Table 5.3–1 presents values for $V_{dc}(t)$ and the peak emitter current $I_P(t)$ vs. t/τ for the expressions given in Eq. (5.3–15). [$I_P(t)$ is the coefficient of $W(t)$ in Eq. (5.3–14) or Eq. (5.3–15a).] In addition, a sketch of $i_E(t)$, given by Eq. (5.3–15a), is shown in Fig. 5.3–3.

It is apparent from Fig. 5.3–3 that at $t = 0$, when $v_i(t)$ adds directly to $V_{dc}(0) = -V_C(0)$ across the input exponential characteristic of the transistor, an exceptionally large emitter current flows with waveform $W(t)$. The average value of this large current, flowing primarily through the coupling capacitor, very rapidly decreases V_{dc} and in turn reduces the peak value of the emitter current. In particular for $t/\tau = 1$

Table 5.3–1 Tabulation of values for $V_{dc}(t)$ and $I_P(t)$ vs. t/τ

t/τ	$I_P(t)$	$V_{dc}(t)$, mV
0	$2200I_{dc}$	760
1	$12.6I_{dc}$	626
2	$9.25I_{dc}$	618
3	$8.43I_{dc}$	614
4	$8.15I_{dc}$	614
5	$8I_{dc}$	614

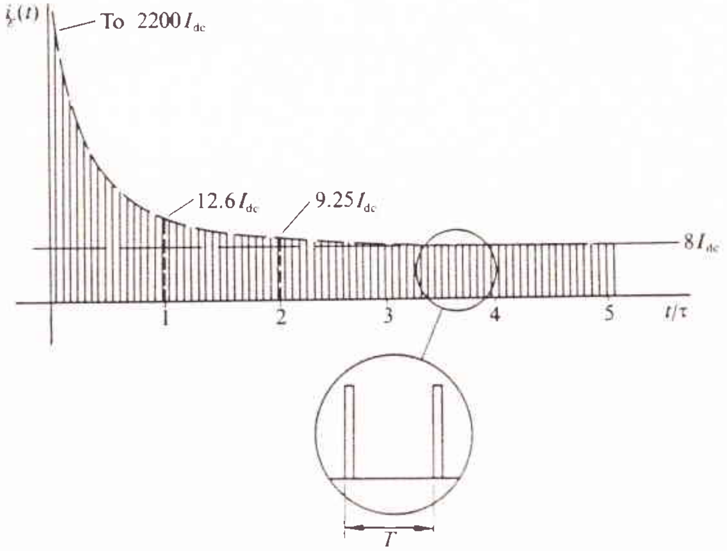


Fig. 5.3–3 Sketch of $i_E(t)$ vs. t for $v_i(t)$ given in Fig. 5.3–2.

the peak value of the emitter current has fallen from $2200I_{dc}$, $(t/\tau = 0)$ to $12.7I_{dc}$. By the time $t/\tau = 4$, steady state has been reached for all practical purposes.

The practical difficulties with this solution are that for any appreciable $v_i(t)$ the turn-on current is so enormous that the bulk of the emitter current transient is over by the time $t = 0.10\tau$ or less. The practical consequences of the demand for this large current are several. First, C must be *very* large or the assumption of a small change of capacitor voltage per cycle is violated. Second, the supposedly very large *instantaneous* currents lead to appreciable voltage drops across internal transistor (or diode) resistances, so that we cease to have the true exponential voltage-current relationship that was assumed. We have instead a resistor-transistor combination of the type assumed in Section 4.8. The actual transient in a real circuit is somewhere between the extreme value suggested by Fig. 5.3–3 and the value obtained from a resistor battery model formed by the techniques of Sections 5.1 and 4.8. The reason is primarily that the presence of the resistance limits the initial emitter current amplitude for $0 \leq t < 0.1\tau$ without appreciably affecting the remainder of the transient.

For the case where $v_i(t) = V_1 \cos \omega t$, $W_x(t)$ has the form shown in Fig. 4.5–3,

$$\overline{W_x} = \frac{I_0(x)}{e^x}, \qquad \text{and} \qquad e^{-qv_{i\,\max}/kT} = e^{-x};$$

hence from Eqs. (5.3–13b) and (5.3–14) we obtain

$$V_{dc}(t) = \frac{kT}{q} \ln \frac{I_{dc}/I_{ES}}{I_0(x) + [1 - I_0(x)]e^{-t/\tau}} \tag{5.3–16}$$

and

$$i_E(t) = \frac{I_{dc}[W_x(t)]e^x}{I_0(x) + [1 - I_0(x)]e^{-t/\tau}}. \tag{5.3–17}$$

A sketch of $i_E(t)$ in this case has essentially the same form as the sketch of Fig. 5.3–3, except that the pulse shape is given by $W_x(t)$. Consequently, here again the emitter current starts with a very large peak value and rapidly decreases toward its steady-state level, which it essentially attains for $t > 4\tau$. Again the internal transistor resistance or the generator impedance limits the initial value of the emitter current.

At this point, it is apparent that the response time of the circuit of Fig. 5.3–1 is governed by the time constant $\tau = C/g_{in}$. Therefore, if $v_i(t)$ is an AM wave and the envelope of this AM wave varies slowly compared with τ, the circuit essentially reaches steady state for each envelope value. Consequently, any steady-state transfer function relating the peak value of $v_i(t)$ to a current or voltage in the circuit of Fig. 5.3–1 can also be used to relate the time-varying envelope of $v_i(t)$ to the corresponding current or voltage.

For example, with $v_1(t) = V_1 \cos \omega t$, the fundamental steady-state collector current is given by

$$i_{C1} = \alpha I_{dc} \frac{2I_1(x)}{I_0(x)} \cos \omega t. \tag{5.3–18}$$

where $x = qV_1/kT$. However, if $v_i(t) = g(t) \cos \omega t$ and $g(t)$ varies slowly in comparison with τ, then

$$i_{C1} = \alpha I_{dc} \frac{2I_1[qg(t)/kT]}{I_0[qg(t)/kT]} \cos \omega t. \tag{5.3–19}$$

Again we note from Fig. 4.5–5 that if $g(t)$ is large in comparison with kT/q for all values of t, Eq. (5.3–19) reduces to

$$i_{C1} \approx 2\alpha I_{dc} \cos \omega t$$

and the circuit strips the AM information from i_{C1}, thus functioning as a limiter.

Note carefully that if the circuit of Fig. 5.3–1 is used to process AM signals, the value of the coupling capacitor C must be arrived at by a compromise. The value must be large enough to appear as a short circuit at the fundamental frequency of $v_i(t)$, and yet it must be small enough so that $\tau = C/g_{in}$ is small in comparison with the time over which the envelope of the AM wave varies. Such a compromise is possible only if there is a wide separation (of a factor of at least 100) between the maximum modulation frequency and the carrier frequency.

In Section 5.2 we saw that the coupling capacitor appears as a short circuit at the frequency ω if

$$\tau = \frac{C}{g_{in}} > 20T,$$

or equivalently,

$$\frac{\omega C}{g_{in}} > 40\pi. \tag{5.3–20}$$

Thus Eq. (5.3–20) puts a limit on the minimum value of C.

To obtain an inequality for the maximum coupling capacitor value when using the circuit of Fig. 5.3–1 to process input AM signals, it would be desirable to obtain a general expression for the capacitor voltage and the emitter current as a function of the envelope of $v_i(t)$ and the capacitor value. With such a general expression, the maximum capacitor value which produces the desired emitter current could readily be determined. Unfortunately, a general relationship is mathematically intractable; however, a small-signal transfer function relating low-index envelope variations to small variations in capacitor voltage and emitter current is possible. In addition, the maximum coupling capacitor value obtained from the small-signal analysis again provides an excellent estimate of the value required for large-signal operation. This small-signal transfer function will be derived in Section 5.6.

5.4 CAPACITIVELY COUPLED TRANSISTOR AMPLIFIER—RESISTOR BIAS

In this section we continue our analysis of the large-signal capacitively coupled transistor amplifier by considering the circuit shown in Fig. 5.4–1, which is biased in a conventional fashion. Here, in addition to a coupling capacitor C_E, a base bypass

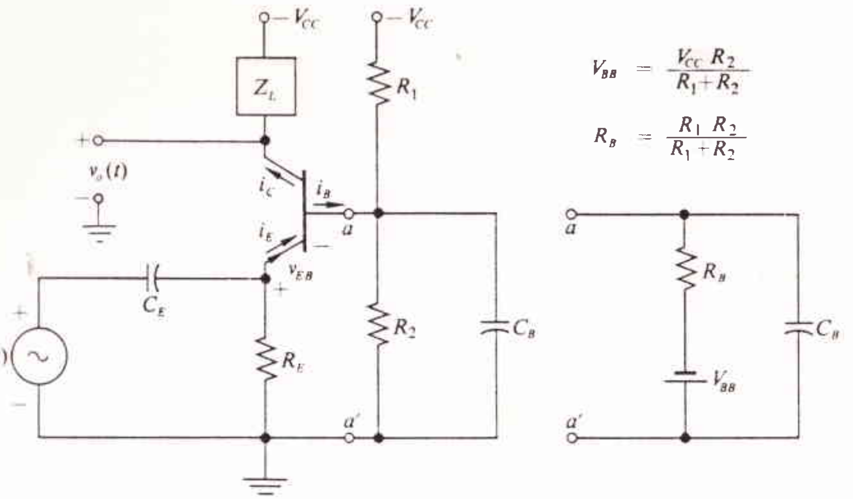


Fig. 5.4–1 Capacitively coupled transistor amplifier with conventional bias.

capacitor is included to ensure that the entire periodic ac signal $v_i(t)$ appears across the emitter-base junction. We again begin by determining the steady-state behavior of the circuit. This analysis is somewhat more involved than the corresponding analysis of the circuit of Fig. 5.3–1, because the depression in the dc emitter-base voltage caused by the application of $v_i(t)$ results in a change in the average emitter current from its quiescent ($v_i = 0$) value. Hence, in addition to the average emitter-base voltage, the average emitter current must be determined as a function of $v_i(t)$ before the total emitter current and in turn the output voltage can be evaluated.

Fortunately, as we shall demonstrate, the change in the average emitter current with input drive level is usually a second-order effect in most well-designed transistor amplifiers: hence I_{EQ} merely replaces I_{dc} in all previously derived results. For the few circuits which are exceptions to the rule, we develop the set of universal curves for $G_m(x)/g_{mQ}$ shown in Fig. 4.5–5 which directly replaces the curve of $G_m(x)/g_m$ of Fig. 4.5–6 in the analysis of narrowband circuits with sinusoidal drives.

We determine the steady-state average emitter current I_{E0} as a function of $v_i(t)$ by first computing the depression ΔV of the emitter-base voltage from its quiescent value when v_i is applied, and then we employ ΔV to compute the shift in I_{E0} from its quiescent value I_{EQ}.

When $x > 4$ for sinusoidal drives, the clamping model of Fig. 5.1–6 may be used in two successive approximations to provide a very rapid answer for this circuit. (See Problem 5.5 at the end of this chapter for an example.)

In the analysis of the circuit of Fig. 5.4–1 we again assume that the transistor does not saturate, that

$$i_C = \alpha i_E \qquad \text{or} \qquad i_B = (1 - \alpha) i_E,$$

that internal transistor reactances are negligible, and that

$$i_E = I_{ES} e^{q v_{EB}/kT},$$

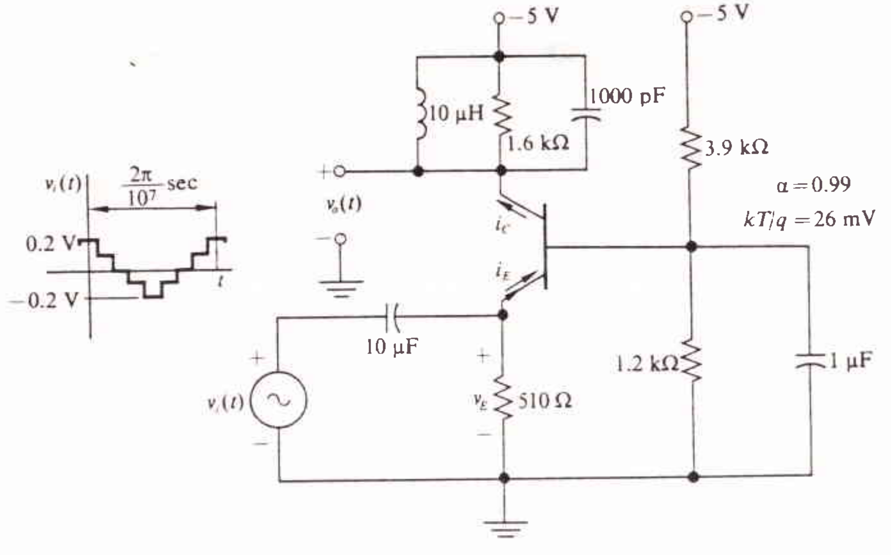


Fig. 5.4–2 Narrowband transistor amplifier.

With these assumptions and the Thévenin equivalent of the base circuit shown in Fig. 5.4–1, the quiescent emitter-base voltage V_{dcQ} and the quiescent emitter current I_{EQ} [which exist prior to applying $v_i(t)$] are related by the equations

$$V_{dcQ} = \frac{kT}{q} \ln \frac{I_{EQ}}{I_{ES}} \tag{5.4–1}$$

and

$$I_{EQ} = \frac{V_{BB} - V_{dcQ}}{R_E + (1 - \alpha) R_B}. \tag{5.4–2}$$

In general, Eqs. (5.4–1) and (5.4–2) are difficult to solve simultaneously; however, if we assume

$$I_{ES} = 2 \times 10^{-16}\ \text{A} \qquad \text{and} \qquad kT/q = 26\ \text{mV}$$

and tabulate V_{dcQ} vs. I_{EQ} as shown in Table 5.4–1, we find that V_{dcQ} experiences only small variations about 760 mV for values of I_{EQ} between 0.1 mA and 10 mA. Therefore, for many practical circuits, V_{dcQ} may be closely approximated by $\frac{3}{4}$ V, and with this value of V_{dcQ}, Eq. (5.4–2) simplifies to

$$I_{EQ} = \frac{V_{BB} - (\frac{3}{4}\ \text{V})}{R_E + (1 - \alpha) R_B}. \tag{5.4–3}$$

The value of $I_{ES} = 2 \times 10^{-16}$ A corresponds to a silicon transistor integrated on a "chip." For a nonintegrated silicon transistor, $I_{ES} \approx 2 \times 10^{-14}$ A and V_{dcQ} may be approximated by 650 mV. For a germanium transistor, $I_{ES} \approx 2 \times 10^{-7}$ A and 220 mV provides a reasonable approximation for V_{dcQ}.

Table 5.4–1 Tabulation of V_{dcQ} vs. I_{EQ} with $kT/q = 26$ mV and $I_{ES} = 2 \times 10^{-16}$ A

V_{dcQ}, mV	I_{EQ}, mA
700	0.1
760	1
820	10
880	100

If we wish the quiescent point to remain fixed, independent of temperature or roduction variations in the transistor parameters, I_{EQ} must be insensitive to these ariations. If I_{EQ} does not vary, $I_{CQ} = \alpha I_{EQ}$ remains fixed, since $\alpha \approx 1$ and in turn $_{CEQ}$, which is a function of I_{EQ}, I_{CQ}, and the bias circuitry, is stable. We observe from q. (5.4–2) that, to keep I_{EQ} independent of transistor parameters, the following onditions must be satisfied:

$$R_E \gg (1-\alpha)R_B \tag{5.4–4a}$$

nd

$$V_{BB} \gg V_{dcQ}. \tag{5.4–4b}$$

The greater the strength of the above inequalities, the greater the bias stability with espect to transistor parameter variations. If $R_E \approx R_B$ and $V_{BB} \approx 10\ V_{dcQ}$, excellent ias stability is achieved in most applications.

If now the periodic input $v_i(t)$ is applied and both C_B and C_E are assumed to be ic short circuits over the frequency range occupied by $v_i(t)$, the emitter current for he circuit of Fig. 5.4–1 is given by

$$\begin{aligned} i_E &= I_{ES}e^{qV_{dc}/kT}e^{qv_i(t)/kT} \\ &= I_{ES}e^{qV_{dc}/kT}e^{qv_{i\max}/kT}W(t), \end{aligned} \tag{5.4–5}$$

vhere V_{dc} is the dc voltage which develops across the combination of C_B and C_E. We observe here that the emitter current waveform is still determined by $e^{qv_i(t)/kT}$ and is therefore invariant with changes in bias configuration. In addition, we observe that, if $-v_i(t)$ were inserted in series with C_B and C_E were returned to ground, Eq. (5.4–5) would remain unchanged; hence, if $v_i(t) = V_1 \cos \omega t$, the only difference between driving the emitter and driving the base would be the shift in the emitter current by half of a carrier cycle.

If we again define the steady-state emitter-base voltage bias depression ΔV as

$$\Delta V = V_{dcQ} - V_{dc}, \tag{5.4–6}$$

then the expression for the emitter current may be rewritten as

$$i_E(t) = I_{EQ}e^{-q\Delta V/kT}e^{qv_{i\max}/kT}W(t). \tag{5.4–7}$$

vhere $I_{EQ} = I_{ES}e^{qV_{dcQ}/kT}$. In addition, the average emitter current I_{E0} is given by

$$I_{E0} = I_{EQ}e^{-q\Delta V/kT}e^{qv_{i\max}/kT}\overline{W(t)}. \tag{5.4–8}$$

In the steady state, I_{E0} is also constrained by the bias circuitry to have the form

$$I_{E0} = \frac{V_{BB} - V_{dc}}{R_E + (1-\alpha)R_B} = I_{EQ} + \frac{\Delta V}{R_E + (1-\alpha)R_B} = I_{EQ}\left(1 + \frac{\Delta V}{V_\lambda}\right), \tag{5.4–9}$$

where $V_\lambda = I_{EQ}[R_E + (1-\alpha)R_B] = V_{BB} - V_{dcQ}$ is the sum of the quiescent voltage which appears across R_B and the quiescent voltage which appears across R_E. Combining Eqs. (5.4–8) and (5.4–9), we obtain an expression for the steady-state bias depression ΔV:

$$\left(1 + \frac{\Delta V}{V_\lambda}\right)e^{q\Delta V/kT} = e^{qv_{i\max}/kT}\overline{W(t)}. \tag{5.4–10}$$

Because of the transcendental nature of the expression for ΔV given in Eq. (5.4–10), it is impossible to obtain an explicit solution for ΔV in closed form. Fortunately, however, for $V_\lambda > 520$ mV [which is always true in a well-designed circuit—cf. Eq. (5.4–4b)], the value of ΔV satisfying the equation

$$e^{q\Delta V/kT} = e^{qv_{i\max}/kT}\overline{W(t)} \tag{5.4–11}$$

differs by less than 5% from the value of ΔV satisfying Eq. (5.4–10), and the agreement is closer with higher values of V_λ.† Consequently, with the approximation for ΔV given in Eq. (5.4–11) we obtain the closed-form expression

$$\Delta V = v_{i\max} + \frac{kT}{q}\ln \overline{W(t)}. \tag{5.4–12}$$

With the aid of Eq. (5.4–9) we also obtain

$$I_{E0} = I_{EQ}\left[1 + \frac{(qv_{i\max}/kT) + \ln \overline{W}}{qV_\lambda/kT}\right]. \tag{5.4–13}$$

By noting [from Eqs. (5.4–7) and (5.4–8)] that $i_E(t)/I_{E0} = W(t)/\overline{W(t)}$, we may finally

---

† Assume that ΔV_1 satisfies Eq. (5.4–11). Therefore, equating Eqs. (5.4–10) and (5.4–11), we obtain

$$\left(1 + \frac{q\Delta V/kT}{qV_\lambda/kT}\right)e^{q\Delta V/kT} = e^{q\Delta V_1/kT},$$

or equivalently,

$$\frac{\Delta V_1 - \Delta V}{\Delta V} = \frac{kT}{q\,\Delta V}\ln\left(1 + \frac{q\Delta V/kT}{qV_\lambda/kT}\right).$$

However, $\ln(1 + x) \le x$; therefore,

$$\frac{\Delta V_1 - \Delta V}{\Delta V} \le \frac{1}{qV_\lambda/kT}$$

and, for $qV_\lambda/kT > 20$,

$$\left|\frac{\Delta V - \Delta V_1}{\Delta V}\right| < 0.05.$$

rite the expression for the total emitter current in the form

$$i_E(t) = I_{EQ}\left[1 + \frac{(qv_{i_{\max}}/kT) + \ln \overline{W}}{qV_\lambda/kT}\right]\frac{W(t)}{\overline{W(t)}} = I_P W(t), \qquad (5.4\text{–}14)$$

here $I_P = I_{E0}/\overline{W}$. Consequently, by employing Eq. (5.4–14), we can determine the teady-state expression for $i_E(t)$ for any bias arrangement similar to the one shown n Fig. 5.4–1, once the form of $v_i(t)$ is known.

As an example of determining the steady-state value of $i_E(t)$ as well as the steady-tate output voltage, we consider the circuit shown in Fig. 5.4–2 with $v_i(t)$ of the form hown in Fig. 5.3–2 applied. We assume that the transistor is silicon and not part of n integrated circuit, so that $V_{dcQ} \approx 0.65$ V, and that $\alpha = 0.99$. From the bias ircuitry we observe that

$$V_{BB} = (5\text{ V})(1.2\text{ k}\Omega)/5.1\text{ k}\Omega = 1.18\text{ V} \qquad \text{and} \qquad R_B = 1.2\text{ k}\Omega \| 3.9\text{ k}\Omega = 916\,\Omega;$$

ence

$$I_{EQ} = \frac{1.18\text{ V} - 0.65\text{ V}}{510\,\Omega + 9.16\,\Omega} = 1.02\text{ mA}$$

nd $V_\lambda = 1.18\text{ V} - 0.65\text{ V} = 0.53$ V. In addition, since $v_{i_{\max}} = 200$ mV, $kT/q =$ 6 mV, and $\overline{W} = \frac{1}{8}$, we obtain from Eq. (5.4–13) $I_{E0} = 1.02\text{ mA}(1.276) = 1.3$ mA, hich leads directly to

$$i_E(t) = 8(1.3\text{ mA})W(t) = (10.4\text{ mA})W(t),$$

here $W(t)$ is the train of narrow pulses shown in Fig. 5.3–2. Since the tuned collector ircuit has $Q_T = 16$ and is resonant at $\omega = 10^7$ rad/sec, the output voltage is given by

$$v_0(t) \approx (\alpha)(2)(1.3\text{ mA})(1.6\text{ k}\Omega)\cos 10^7 t - (5\text{ V})$$
$$= (4.16\text{ V})\cos 10^7 t - (5\text{ V}).$$

ere again, since the emitter current pulses are narrow, we have approximated the mplitude of the fundamental component as twice the average value.

In solving this problem we have assumed that the transistor does not saturate or, quivalently, that the collector voltage remains more negative than the emitter oltage at all times. To check this assumption we require, in addition to v_0, an xpression for the emitter voltage v_E. Clearly, v_E contains the dc component $-I_{E0}R_E$ nd the ac component $v_i(t)$; therefore,

$$v_E(t) = -(0.662\text{ V}) + v_i(t).$$

igure 5.4–3 shows a sketch of $v_E(t)$ and v_0 on the same set of coordinates. From this ketch, which also includes $i_E(t)$, it is apparent that $v_0 < v_E$ at all times and saturation oes not occur.

As a second example, we consider $v_i(t) = V_1 \cos \omega t$ applied to the circuit of Fig. .4–1. For this drive the bias depression ΔV obtained from Eq. (5.4–12) is

$$\Delta V = \frac{kT}{q} \ln I_0(x).$$

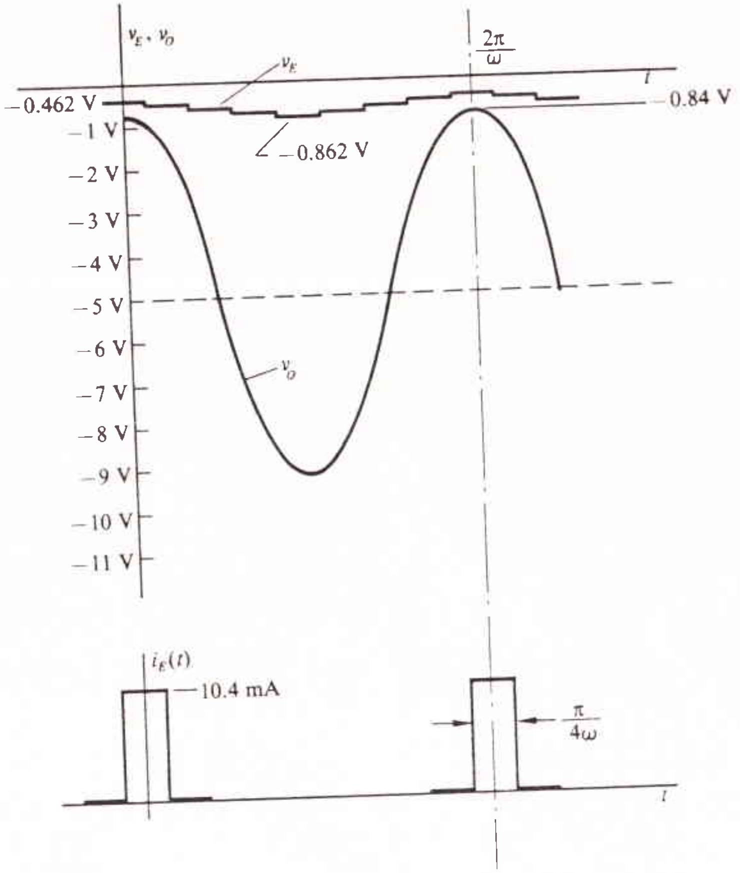


Fig. 5.4–3 Waveforms for the circuit of Fig. 5.4–2.

Table 5.4–2

x	$\ln I_0(x)$	x	$\ln I_0(x)$
0	0	8	6.058
1	0.2359	9	6.998
2	0.8242	10	7.943
3	1.585	12	9.850
4	2.425	14	11.77
5	3.305	16	13.70
6	4.208	18	15.64
7	5.128	20	17.59

Table 5.4–2 includes values of $q\,\Delta V/kT = \ln I_0(x)$ vs. x for values of x between 0 and 20. In addition, with $v_i(t) = V_1 \cos \omega t$, the average emitter current reduces to

$$I_{E0} = I_{EQ}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right]; \qquad (5.4\text{–}15)$$

as a result, the total emitter current is given by

$$\begin{aligned} i_E(t) &= I_{EQ}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right]\frac{e^{x\cos\omega t}}{I_0(x)} \\ &= I_{EQ}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right]\left[1 + 2\sum_{n=1}^{\infty}\frac{I_n(x)}{I_0(x)}\cos n\omega t\right]. \qquad (5.4\text{–}16) \end{aligned}$$

A plot of I_{E0}/I_{EQ} vs. x obtained from Eq. (5.4–15) with qV_λ/kT as a parameter is given in Fig. 5.4–4. With the aid of Fig. 5.4–4, once the quiescent values of I_{EQ} and V_λ and the input level V_1 (or x) are known, the total emitter current or any of its harmonics can be determined.

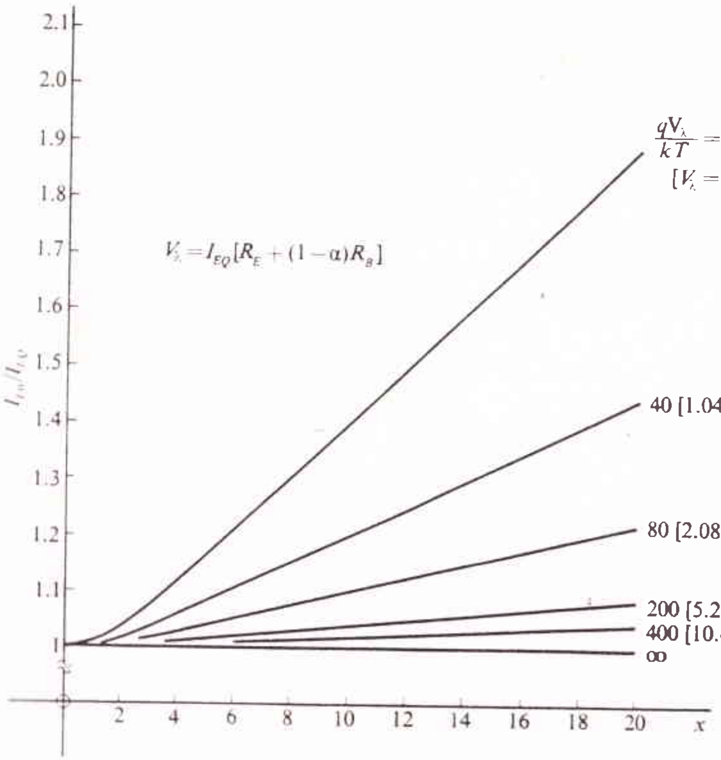


Fig. 5.4–4 Plot of I_{E0}/I_{EQ} vs. x with qV_λ/kT as a parameter. Values of V_λ shown in brackets are obtained with $kT/q = 26$ mV.

For example, consider $v_i = (208 \text{ mV}) \cos 10^7 t$ applied to the circuit of Fig. 5.4–2. Since for this circuit $I_{EQ} = 1.02$ mA, $V_\lambda = 0.53$ V, and $V_1 = 208$ mV, if $kT/q = 26$ mV, then $x = 8$ and $qV_\lambda/kT \approx 20$. From the curves of Fig. 5.4–4 we obtain

$$\frac{I_{E0}}{I_{EQ}} = 1.3$$

and in turn

$$\begin{aligned} i_E(t) &= \frac{(1.33 \text{ mA})e^{8\cos\omega t}}{I_0(8)} \\ &= (9.24 \text{ mA})W_8(t). \end{aligned}$$

In addition, the fundamental component of $i_E(t)$ has the value

$$i_{E1}(t) = (1.33 \text{ mA})\underbrace{\frac{2I_1(8)}{I_0(8)}}_{1.871}\cos\omega t$$

and thus $v_o(t) = (4.0 \text{ V}) \cos \omega t - (5.0 \text{ V})$.

Note carefully that Fig. 5.4–4 provides a set of universal curves that can be used when $v_i = V_1 \cos \omega t$ for any bias circuit values and for either germanium or silicon transistors. However, when dealing with narrowband amplifiers and oscillators, it is more convenient to develop a similar set of universal curves for the large-signal average fundamental transconductance $G_m(x)$, where

$$G_m(x) = \frac{\alpha I_{E1}}{V_1} = \frac{I_{C1}}{V_1}$$

and I_{E1} and I_{C1} are the coefficients of the fundamental components of the emitter and collector currents respectively. With the aid of Eq. (5.4–16) we can express $G_m(x)$ in the form

$$\begin{aligned} G_m(x) &= \frac{\alpha I_{EQ}}{V_1}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right]\frac{2I_1(x)}{I_0(x)} \\ &= g_{mQ}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right]\frac{2I_1(x)}{xI_0(x)}, \qquad (5.4\text{–}17) \end{aligned}$$

where $g_{mQ} = \alpha I_{EQ}/kT$ is the small-signal transconductance evaluated at the quiescent point. The normalized curves of $G_m(x)/g_{mQ}$ vs. x with qV_λ/kT as a parameter are given in Fig. 5.4–5.

Once $G_m(x)$ is determined for any particular set of bias values and input voltage levels from the universal curves of Fig. 5.4–5, the output voltage v_0 of a narrowband amplifier may be expressed directly as

$$v_o(t) = (V_1 \cos \omega t)G_m(x)R_L - V_{CC}, \qquad (5.4\text{–}18)$$

where R_L is the resistance of the collector-tuned circuit at the resonant frequency ω.

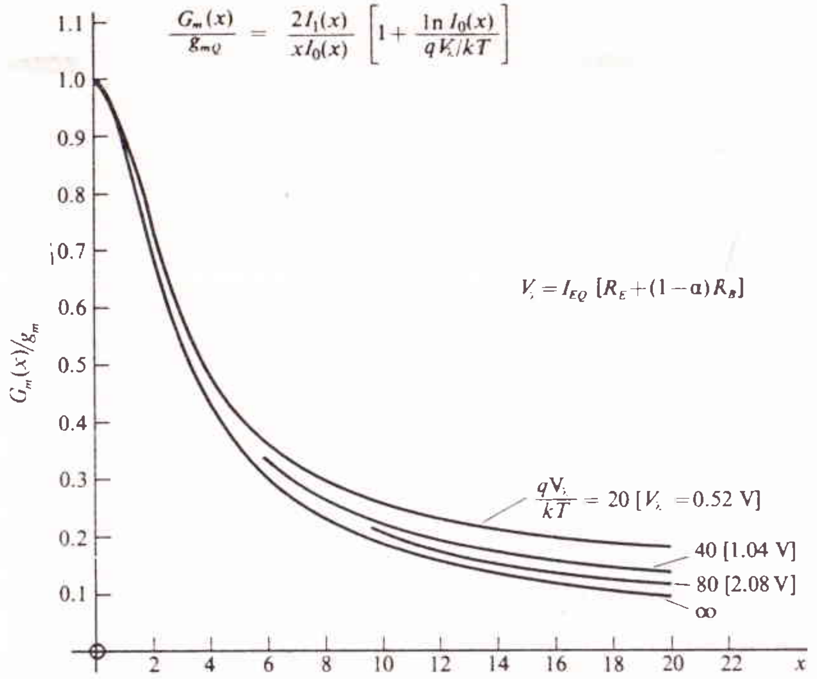


Fig. 5.4–5 Plot of $G_m(x)/g_{mQ}$ vs. x with qV_λ/kT as a parameter. Values of V_λ in brackets are obtained with $kT/q = 26$ mV.

For example, if we again consider $v_i(t) = (208\ \mathrm{mV})\cos 10^7 t$ applied to the circuit of Fig. 5.4–2, then $V_\lambda = 530$ mV and $I_{EQ} = 1.02$ mA. If we let $kT/q = 26$ mV, then $x = 8$, $qV_\lambda/kT \approx 20$, and $g_{mQ} = qI_{EQ}/kT = 39.2$ mmho. From Fig. 5.4–5 we obtain $G_m(x)/g_{mQ} = 0.307$; hence $G_m(x) = 12$ mmho. Finally, since $R_L = 1.6\ \mathrm{k\Omega}$, $v_o(t)$ is given by

$$v_o(t) = -(5\ \mathrm{V}) + (208\ \mathrm{mV})(\cos\omega t)(12\ \mathrm{mmho})(1.6\ \mathrm{k\Omega}) = -(5\ \mathrm{V}) + (4\ \mathrm{V})\cos\omega t,$$

which is in exact agreement with the expression derived by the previous method.

In most well-designed circuits of the form shown in Fig. 5.4–1, $v_{i_{\max}}$ is much less than V_λ or, equivalently, V_{BB} (cf. Eq. 5.4–4b). For this type of circuit,

$$\frac{(qv_{i_{\max}}/kT) + \ln \overline{W}}{qV_\lambda/kT} \ll 1 \tag{5.4–19}$$

(since $\ln \overline{W}$ must be negative and smaller in magnitude than $qv_{i_{\max}}/kT$, the numerator of Eq. (5.4–19) is always less than $qv_{i_{\max}}/kT$) and Eq. (5.4–13) for the average steady-state emitter current simplifies to

$$I_{E0} = I_{EQ}. \tag{5.4–20}$$

Equation (5.4–20) is intuitively satisfying, since a small ac input voltage can at most effect a small dc bias depression, this, when compared with a large V_B, has little effect on the average emitter current. In such a circuit the emitter resistor R_E can be replaced by a current source I_{EQ} with no effect on any of the current or voltage waveshapes within the entire circuit; however, with such a current source bias, all the results of the previous section, as well as all the results of Section 4.5, are directly applicable.

For the case where $v_i(t) = V_1 \cos\omega t$, the inequality of Eq. (5.4–19) is equivalent to (cf. Eq. 5.4–16)

$$\frac{\ln I_0(x)}{qV_\lambda/kT} \ll 1. \tag{5.4–21}$$

Clearly, then, if

$$V_\lambda \geq 20\frac{kT}{q}\ln I_0(x) \tag{5.4–22}$$

for a particular drive level and bias configuration, I_{E0} may be approximated by I_{EQ} within 5%. Figure 5.4–6 plots V_λ, given by an equality in Eq. (5.4–22), vs. x. Thus for any circuit biased as shown in Fig. 5.4–1 having values of V_λ and x which determine

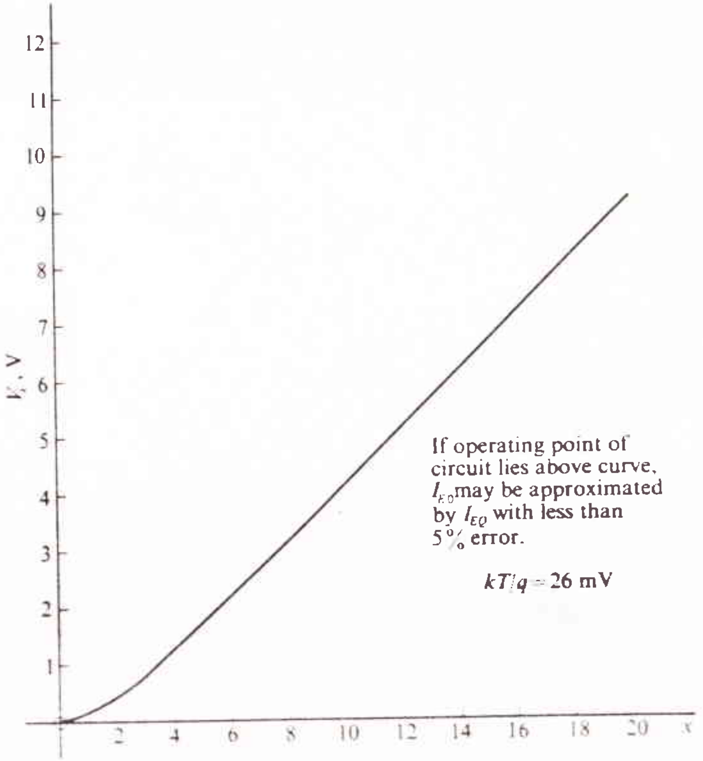


Fig. 5.4–6 Plot of V_λ vs. x for the condition where $I_{E0}/I_{EQ} = 1.05$ or $V_\lambda = 20(kT/q)\ln I_0(x)$.

a point above the curve of Fig. 5.4–6, I_{E0} may be approximated by I_{EQ} with less than 5% error. In addition, $G_m(x)/g_{mQ}$ may be approximated by $2I_1(x)/xI_0(x)$ with less than 5% error. This approximation corresponds to employing the $qV_\lambda/kT = \infty$ curve of Fig. 5.4–5 or, equivalently, the curve of Fig. 4.5–6 to determine $G_m(x)/g_{mQ}$ for any given value of x. It turns out in practice that the operating points of most circuits with stable bias configurations and reasonable drive levels are above the curve of Fig. 5.4–6.

Example 5.4–1 For the circuit shown in Fig. 5.4–7, determine an expression for $v_0(t)$.

Solution. The first thing to note is that we have an NPN transistor rather than a PNP transistor and that we are driving the base and not the emitter. However, since the v_{BE}-i_E characteristic for an NPN transistor is identical with the v_{EB}-i_E characteristic for a PNP transistor, driving the base of an NPN transistor has the same effect on the emitter current as driving the emitter of a PNP transistor. The main difference between the two cases is that the quiescent value of V_{EB} experiences a change in sign.

At quiescence ($v_i = 0$), if we assume that the transistor is silicon ($V_{dcQ} = 0.65$ V),

$$I_{EQ} = \frac{9.35\text{ V}}{5.1\text{ k}\Omega} = 1.83\text{ mA} \qquad \text{and} \qquad V_\lambda = 9.35\text{ V}.$$

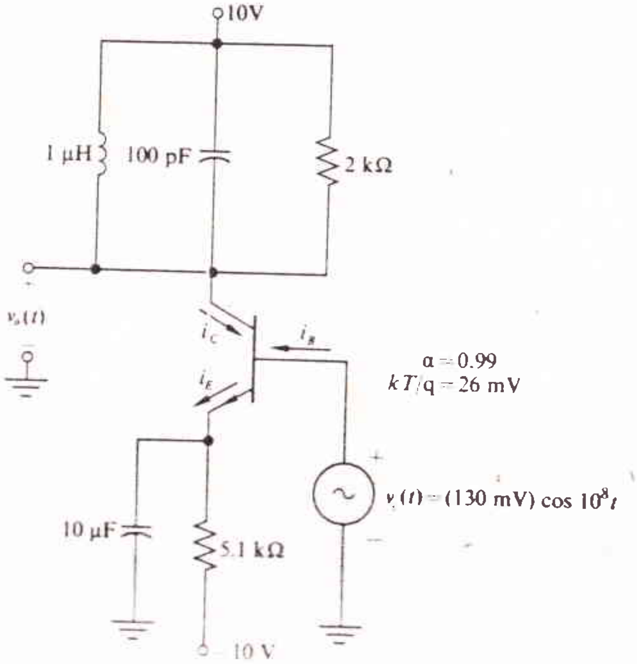


Figure 5.4–7

Since for this example $x = 5$, the operating point (x, V_λ) falls well above the curve of Fig. 5.4–6, and thus $I_{E0} \approx I_{EQ}$.

Since the tuned circuit in the collector is resonant at $\omega = 10^8$ rad/sec, which is the fundamental frequency component of $i_E(t)$, and since for the tuned circuit $Q_T = 20$, only the fundamental component of $i_E(t)$ contributes appreciably to $v_o(t)$. Consequently, v_o may be expressed as (cf. Eq. 5.3–12)

$$\begin{aligned} v_o(t) &= +V_{CC} - \frac{2I_1(x)}{I_0(x)} I_{EQ} R_L \cos \omega t \\ &= (10\text{ V}) - 1.787(1.83\text{ mA})(2\text{ k}\Omega) \cos 10^8 t \\ &= (10\text{ V}) - (6.5\text{ V}) \cos 10^8 t. \end{aligned}$$

Alternatively we can determine $g_{mQ} = 70.4$ mmho, and from the $V_\lambda = \infty$ curve of Fig. 5.4–5 obtain $G_m(5)/g_{mQ} = 0.359$, from which it follows that $G_m(5) = 25.2$ mmho. Therefore,

$$\begin{aligned} v_o(t) &= +V_{CC} - G_m(x) R_L V_1 \cos \omega t \\ &= (10\text{ V}) - (25.2\text{ mmho})(2\text{ k}\Omega)(130\text{ mV}) \cos 10^8 t \\ &= (10\text{ V}) - (6.5\text{ V}) \cos 10^8 t \end{aligned}$$

There is a minus sign in the expression for $v_o(t)$ because the collector current flows downward through the tuned circuit.

5.5 NONLINEAR LOADING OF TUNED CIRCUITS

In this section we shift our attention to the properties of a parallel *RLC* circuit which is driven by a periodic current source $i_i(t)$ and which is shunted by a nonlinear nonreactive load as shown in Fig. 5.5–1. The current source $i_i(t)$ is assumed to have a period $T = 2\pi/\omega_0$, where ω_0 is the resonant frequency of the tuned circuit; and the loading, although nonlinear, is assumed to maintain the Q of the circuit reasonably high. With these assumptions, we first show that in the steady state the output voltage $v_o(t)$ remains approximately sinusoidal despite the presence of the nonlinear load and that the nonlinear load may be modeled by an equivalent linear resistance.

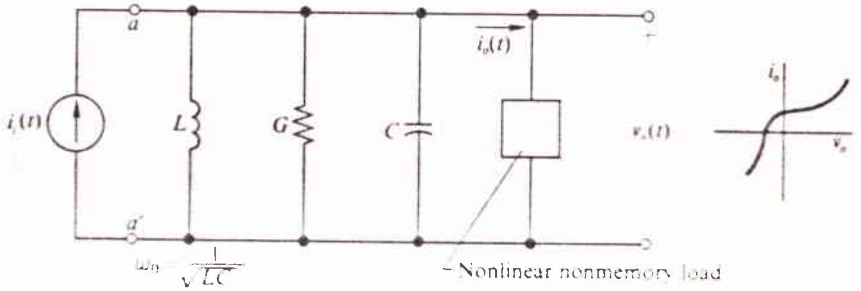


Fig. 5.5–1 Nonlinear loading of parallel *RLC* circuit.

We then develop a graphical technique for determining the steady-state amplitude of $v_o(t)$ as a direct function of the fundamental component of $i_i(t)$.

Circuits of the type shown in Fig. 5.5–1 arise when the output of one large-signal narrowband amplifier (or limiter) drives the nonlinear input characteristic of a subsequent large-signal amplifier, or in the case of an oscillator, when the output of a narrowband amplifier drives its own nonlinear input characteristic. Clearly the determination of the relationship between $v_o(t)$ and $i_i(t)$ is a necessary step in the calculation of the overall amplification of a chain of narrowband amplifiers or of the loop gain of a sine-wave oscillator.

To begin our analyses, we observe that since $i_i(t)$ is periodic it may be expanded in a Fourier series of the form

$$i_i(t) = I_{i0} + I_{i1}\cos\omega_0 t + \sum_{n=2}^{\infty} I_{in}\cos(n\omega_0 t + \theta_{in}), \tag{5.5–1}$$

where the time axis has been chosen in such a fashion that the phase angle of the fundamental component of $i_i(t)$ is equal to zero. In addition, because of the periodicity of $i_i(t)$, in the steady state all other currents and voltages in the circuit must be periodic with period T; in particular, $i_o(t)$ must be periodic and therefore expressible in the form

$$i_o(t) = \sum_{n=0}^{\infty} I_{on}\cos(n\omega_0 t + \theta_{on}). \tag{5.5–2}$$

Without affecting the output voltage $v_o(t)$, the nonlinear load may be replaced by the infinite array of harmonically related current sources contained in $i_o(t)$, as shown in Fig. 5.5–2.

Now the circuit is assumed to have a high value of Q_T, where for the nonlinear circuit Q_T is defined as the ratio of the total resistance R_T seen looking into terminals a-a' (cf. Fig. 5.5–1) at the resonant frequency ω_0 to the reactance of L or C evaluated at ω_0, that is,†

$$Q_T \equiv \frac{R_T}{|X_L(\omega_0)|} = \frac{R_T}{|X_C(\omega_0)|}. \tag{5.5–3}$$

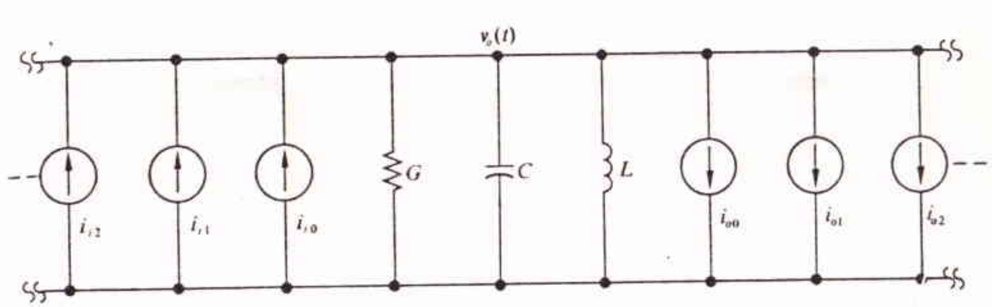
Therefore, the resonant circuit presents a low impedance [less than $|X_C(\omega_0)|$] at dc and all harmonics of ω_0. If, in addition, $|i_{in} - i_{on}|_{\max} \leq I_{i1}$ for all $n \neq 1$ (a condition which is satisfied by almost all practical periodic waveforms), then the dc and harmonic components of $i_i(t)$ and $i_o(t)$ are essentially shunted to ground and contribute very little to the waveform of $v_o(t)$. Thus $v_o(t)$ may be closely approximated by a sinusoid of the form

$$v_o(t) = V_1\cos(\omega_0 t + \theta_{o1}). \tag{5.5–4}$$

The phase angle of $v_o(t)$ must of course be identical with the phase angle θ_{o1} of the fundamental component of $i_o(t)$, since the nonlinear device possesses no energy storage. (The reader should convince himself of this fact.)

---

† This definition is clearly consistent with the definition given in Section 2.2.



$$i_{i1} = I_{i1}\cos\omega_0 t$$
$$i_{in} = I_{in}\cos(n\omega_0 t + \theta_{in})$$
$$i_{on} = I_{on}\cos(n\omega_0 t + \theta_{on})$$

Fig. 5.5–2 Current-source representation of non-linear load.

A simplified circuit for determining $v_o(t)$, from which the dc and harmonic components of $i_i(t)$ and $i_o(t)$ have been omitted, is shown in Fig. 5.5–3.

The problem in determining $v_o(t)$ in terms of $i_i(t)$ is now reduced to evaluating V_1 and θ_{o1}. At resonance, however, it is apparent from Fig. 5.5–3 that

$$I_{i1}\cos\omega_0 t = (V_1 G + I_{o1})\cos(\omega_0 t + \theta_{o1}); \tag{5.5–5}$$

hence $\theta_{o1} = 0$ and

$$I_{o1} = I_{i1} - V_1 G. \tag{5.5–6}$$

To evaluate V_1 we employ a method similar to the one used in Section 5.1 to determine V_C. We assume a value for V_1 which, of course, specifies $v_o(t) = V_1\cos\omega_0 t$. With this value of $v_o(t)$ applied across the nonlinear load, we determine $i_o(t)$ and, either analytically or graphically, determine the amplitude of its fundamental coefficient I_{o1}. This procedure is repeated for a sufficient number of different values of V_1 until a smooth curve of I_{o1} vs. V_1 is obtained. This curve represents the relationship between V_1 and I_{o1} required by the nonlinear load. The linear portion of the circuit, however, requires that V_1 and I_{o1} be related by Eq. (5.5–6). Therefore, if the "load line" of Eq. (5.5–6) is plotted on the same set of coordinates as the V_1-I_{o1} curve, as shown in Fig. 5.5–4, the intersection of the "load line" and the curve determines the steady-state values of V_1 and I_{o1}. With V_1 known, $i_o(t)$ can readily be computed.

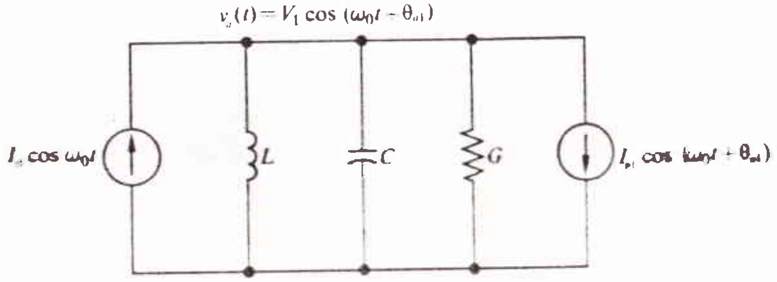

$v_o(t) = V_1\cos(\omega_0 t - \theta_{o1})$
$I_{i1}\cos\omega_0 t$
L
C
G
$I_{o1}\cos(\omega_0 t + \theta_{o1})$

Fig. 5.5–3 Simplified network for determining $v_o(t)$.

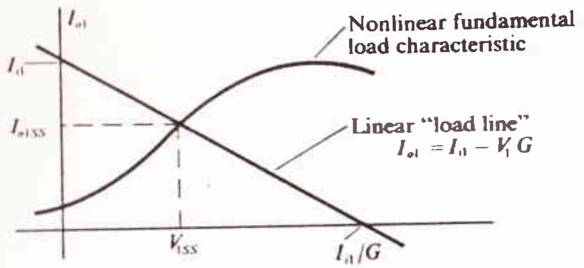


Fig. 5.5–4 Graphical method for determining the steady-state values of V_1 and I_{o1}.

It is desirable to write an expression for Q_T in terms of the steady-state values f V_1 (V_{1SS}) and I_{o1} (I_{o1SS}) such that, once V_{1SS} is found, the assumption of high Q_T an be checked. To accomplish this goal, the model of Fig. 5.5–3 is further simplified y replacing the output current source by an equivalent linear conductance,

$$G_{NL} = \frac{I_{o1}}{V_1},$$

which maintains the same currents and voltages throughout the network. This implified model is shown in Fig. 5.5–5, from which it is immediately apparent that

$$Q_T = \frac{\omega_0 C}{G + G_{NLSS}} = \frac{1}{\omega_0 L(G + G_{NLSS})}, \tag{5.5–7}$$

where

$$G_{NLSS} = \frac{I_{o1SS}}{V_{1SS}}.$$

More important, however, is the fact that nonlinear loading across a high-Q tuned circuit has the same effect as linear resistive loading. We now demonstrate the above method with several examples.

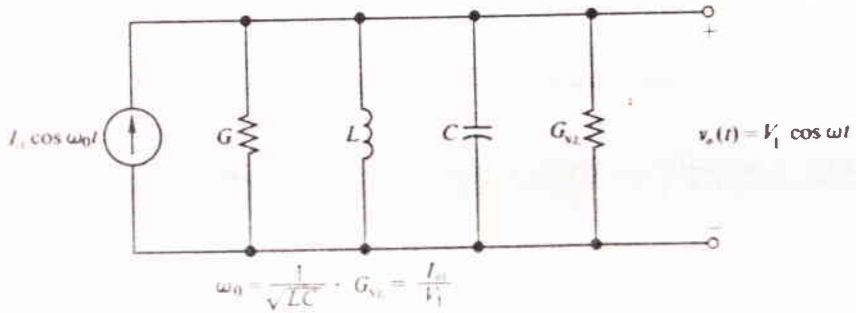

$I_o \cos \omega_0 t$
G
L
C
G_{NL}
+
$v_o(t) = V_1 \cos \omega t$
−
$\omega_0 = \frac{1}{\sqrt{LC}}$; $G_{NL} = \frac{I_{o1}}{V_1}$

Fig. 5.5–5 Simplified model with nonlinear load replaced by equivalent linear conductance.

Example 5.5–1 For the circuit shown in Fig. 5.5–6, find $v_o(t)$.

Solution. We assume $Q_T \geq 10$ and write $v_o(t) = V_1 \cos 10^7 t$. Applying this voltage to the nonlinear load, we observe that $i_o(t)$ is a square wave, with a peak-to-peak amplitude of π, which may be expanded in a Fourier series of the form

$$i_o(t) = (2\text{ mA})(\cos 10^7 t - \tfrac{1}{3}\cos 3 \times 10^7 t + \tfrac{1}{5}\cos 5 \times 10^7 t - \cdots).$$

The resonant circuit, with a high value of Q_T, appears as an effective short circuit to the components of $i_o(t)$ at $\omega = 3 \times 10^7$ rad/sec, 5×10^7 rad/sec, etc; hence, $v_o(t)$ is obtained by multiplying the input current, less the fundamental component of $i_o(t)$, by the tuned-circuit impedance evaluated at $\omega = 10^7$ rad/sec, that is,

$$v_o(t) = (4\text{–}2\text{ mA})(2\text{ k}\Omega)\cos 10^7 t = (4\text{ V})\cos 10^7 t.$$

It is apparent that the same value of $v_o(t)$ would be obtained by replacing the nonlinear load by a resistor of value

$$\frac{1}{G_{NLSS}} = \frac{V_{1SS}}{I_{o1SS}} = \frac{4\text{ V}}{2\text{ mA}} = 2\text{ k}\Omega$$

Consequently,

$$Q_T = \frac{2\text{ k}\Omega \| 2\text{ k}\Omega}{(10^7\text{ rad/sec})(5\ \mu\text{H})} = 20$$

and our assumption is justified.

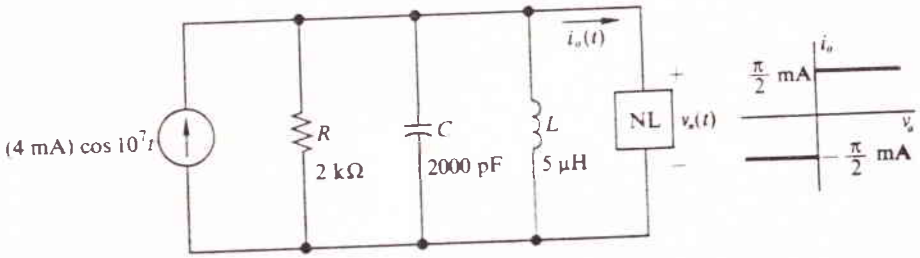


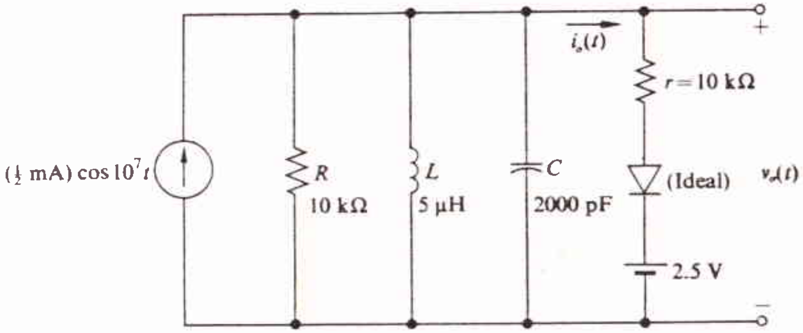
Figure 5.5–6

Example 5.5–2 For the circuit shown in Fig. 5.5–7, determine an expression for $v_o(t)$.

Solution. We again assume $Q_T > 10$ and write $v_o(t) = V_1 \cos 10^7 t$. Applying this voltage to the series diode-battery-resistor combination, we observe that for $V_1 \leq 2.5$ V, $i_o(t) \equiv 0$; thus $I_{o1} = 0$. For $V_1 > 2.5$ V, the waveform of $i_o(t)$ is a periodic train of sine-wave tip pulses of peak amplitude $I_p = (V_1 - V_x)/R$ and conduction angle $2\phi = 2\cos^{-1}(V_x/V_1)$, where $V_x = 2.5$ V and $R = 10$ kΩ. Table 5.5–1 includes values of I_p and I_{o1} for values of V_1 between 2.5 V and 5 V (cf. Fig. 4.2–4), and Fig. 5.5–8 includes a plot of I_{o1} vs. V_1.

Table 5.5–1

V_1, V	I_p, μA	I_{o1}, μA
2.5	0	0
3.0	50	12
3.5	100	30
4.0	150	52
4.5	200	75
5.0	250	98



$$\frac{1}{\sqrt{LC}} = 10^7 \text{ rad/sec}$$

Figure 5.5–7

In addition, since Kirchhoff's current law must hold for currents at the resonant frequency ω_0 flowing into the tuned circuit of Fig. 5.5–7, we obtain the load line

$$I_{o1} = (\tfrac{1}{2}\text{ mA}) - \frac{V_1}{10\ \mathrm{k}\Omega},$$

which is also plotted in Fig. 5.5–8. The two curves intersect at $I_{o1SS} = 67\ \mu\text{A}$ and $V_{1SS} = 4.3$ V. Consequently,

$$\frac{1}{G_{NLSS}} = 65\ \mathrm{k}\Omega \qquad Q_T = \frac{65\ \mathrm{k}\Omega \| 10\ \mathrm{k}\Omega}{(10^7\text{ rad/sec})(5\ \mu\text{H})} = 174$$

(which justifies the high-Q assumption), and finally,

$$v_o(t) = (4.3\text{ V})\cos 10^7 t.$$

Example 5.5–3 For the circuit shown in Fig. 5.5–9, calculate $v_o(t)$.

Solution. We note that this example is identical to the previous example except that the resistance in series with the diode is equal to zero. Clearly, then, when making a table for I_{o1} similar to Table 5.5–1, we find that $I_{o1} = \infty$ for all values of $V_1 > 2.5$ V. Consequently the curve for I_{o1} vs. V_1 takes the form shown in Fig. 5.5–10.

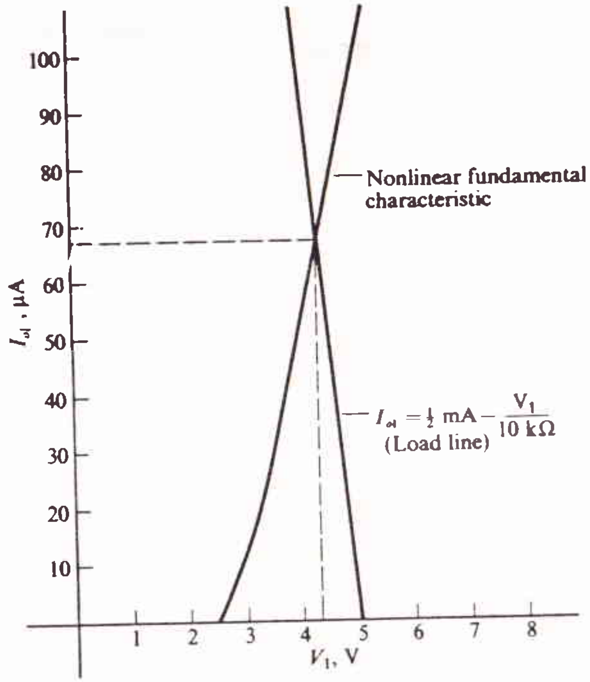


Figure 5.5–8

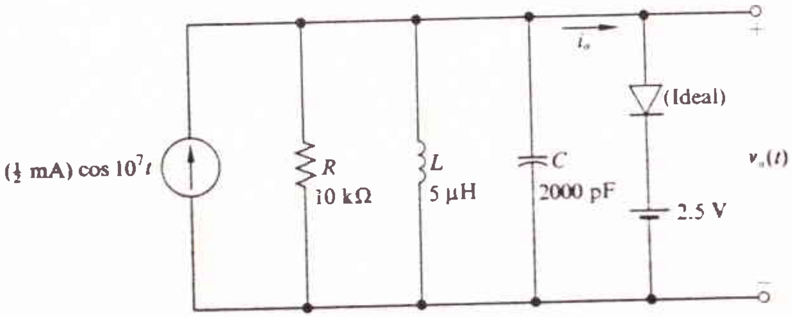


Figure 5.5–9

Again plotting

$$I_{o1} = (\tfrac{1}{2}\text{ mA}) - \frac{V_1}{10\ \mathrm{k}\Omega}$$

on the same set of coordinates, we obtain

$$V_{1SS} = 2.5\text{ V}, \qquad I_{o1SS} = 250\ \mu\text{A}, \qquad \frac{1}{G_{NLSS}} = 10\ \mathrm{k}\Omega$$

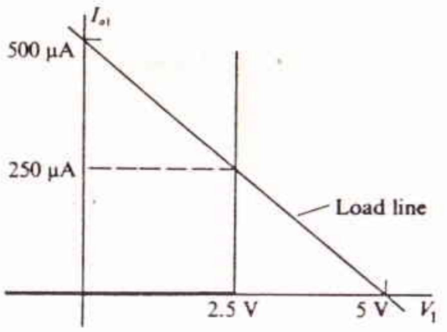


Figure 5.5–10

and

$$Q_T = \frac{10\text{ k}\Omega \| 10\text{ k}\Omega}{(10^7\text{ rad/sec})(5\ \mu\text{H})} = 100.$$

Hence the method of solution is justified.

Note that in this problem, although a battery and diode are placed directly across the tuned circuit, the voltage remains sinusoidal of the form $v_o = (2.5\text{ V})\cos 10^7 t$; hence the major effect of the diode is to clamp the peak-to-peak amplitude of the resonant circuit voltage to twice the battery value. The diode does not, as is sometimes erroneously believed, clip the top of $v_o(t)$ and leave the bottom unchanged. Such a waveform obviously cannot exist, since its average value would be different from zero and thus could not be sustained across an inductor.

The circuit of this example finds application as an FM limiter. Clearly, variations of the input current about the $\frac{1}{2}$ mA point result in the same peak output voltage of $V_1 = 2.5$ V. Only when the input current drops below $\frac{1}{4}$ mA does V_1 drop below its 2.5 V level.

Example 5.5–4 For the circuit shown in Fig. 5.5–11, assuming that no collector saturation occurs and that both input and output tuned circuits have high values of Q_T and both resonate at ω_0, determine an expression for $v_1(t)$ and $v_o(t)$. Also determine the effective linear loading the transistor applies to the input parallel resonant circuit.

Solution. Since Q_T for the input tuned circuit is assumed to be high,

$$v_1(t) = V_1 \cos \omega_0 t \qquad \text{and} \qquad v_{BE} = V_1 \cos \omega_0 t + V_{dc}.$$

Thus if an exponential relationship is assumed to exist between i_E and v_{BE}, $i_E(t)$ is given by

$$i_E(t) = I_{ES} e^{qV_{dc}/kT} e^{x \cos \omega_0 t}$$

and in the steady state the amplitude of the fundamental component of emitter current I_{E1} is given by

$$I_{E1} = I_{dc} \frac{2I_1(x)}{I_0(x)}.$$

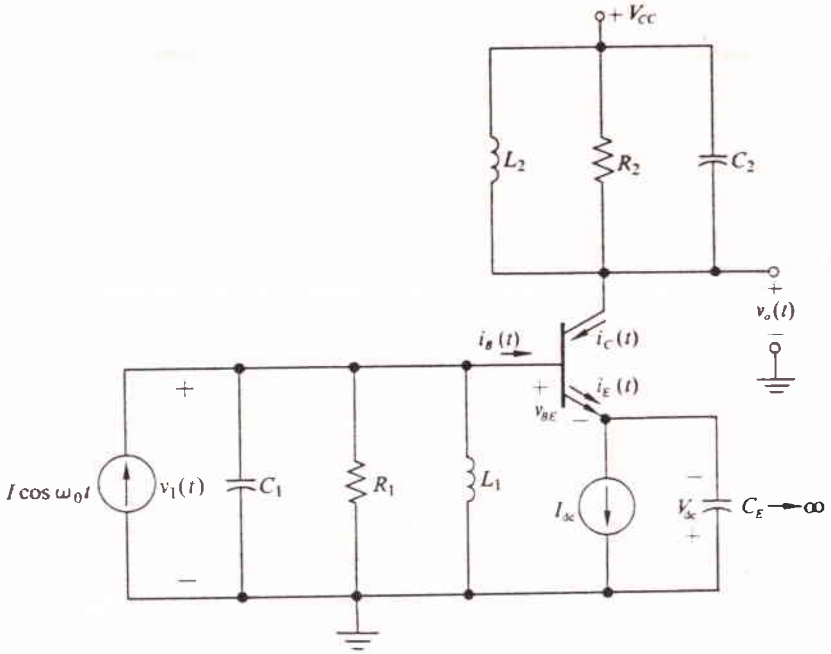


Figure 5.5–11

Since $i_C = \alpha i_E$ and $i_B = (1 - \alpha) i_E$, the amplitudes of the fundamental components of the collector and base currents are given by

$$I_{C1} = \alpha I_{dc} \frac{2I_1(x)}{I_0(x)} \qquad \text{and} \qquad I_{B1} = (1 - \alpha) I_{dc} \frac{2I_1(x)}{I_0(x)}$$

respectively.

Now to obtain an expression for V_1 (or x) we equate the fundamental components of the currents entering the input-tuned circuit to obtain

$$I - \frac{V_1}{R_1} = I_{dc}(1 - \alpha)\frac{2I_1(x)}{I_0(x)},$$

or equivalently,

$$\frac{I}{(1-\alpha)I_{dc}} - \frac{V_1}{(1-\alpha)I_{dc}R_1} = \frac{I}{(1-\alpha)I_{dc}} - \frac{x}{(1-\alpha)g_{m}R_1} = \frac{2I_1(x)}{I_0(x)},$$

where $g_m = qI_{dc}/kT$. By plotting the "load line"

$$\frac{I}{(1-\alpha)I_{dc}} - \frac{x}{(1-\alpha)g_{m}R_1}$$

on the same set of coordinates as $2I_1(x)/I_0(x)$ vs. x, as shown in Fig. 5.5–12, we obtain x_{SS} and $2I_1(x_{SS})/I_0(x_{SS})$ as the coordinates of the intersection of the two curves; thus

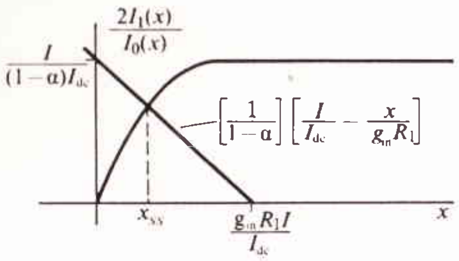


Figure 5.5–12

$$v_1 = \frac{kTx_{SS}}{q}\cos\omega_0 t.$$

Specifically, if $I/(1-\alpha)I_{dc} = 2$ and $(1-\alpha)g_{in}R_1 = 6$, by plotting $2 - x/6$ on the coordinates of Fig. 4.5–5 we obtain

$$x_{SS} = 2, \qquad \frac{2I_1(2)}{I_0(2)} = 1.35, \qquad \text{and} \qquad v_1(t) = (52\text{ mV})\cos\omega_0 t.$$

With $2I_1(x_{SS})/I_0(x_{SS})$ known and the output-tuned circuit Q assumed to be high, $v_o(t)$ can immediately be written as

$$\begin{aligned} v_o(t) &= V_{CC} - I_{C1}R_2\cos\omega_0 t \\ &= V_{CC} - \alpha I_{dc}\frac{2I_1(x_{SS})}{I_0(x_{SS})}R_2\cos\omega_0 t. \end{aligned}$$

The equivalent linear loading on the input-tuned circuit is given by

$$G_{NL} = \frac{I_{B1}}{V_1} = \frac{I_{C1}}{\beta V_1} = \frac{G_m(x)}{\beta}, \tag{5.5–8}$$

where $\beta = \alpha/(1-\alpha)$ and G_m is the large-signal average fundamental transconductance of the transistor, which is plotted as a function of x and normalized to g_m in Fig. 4.5–6. For this bias configuration, $g_m = \alpha I_{dc}q/kT$; G_{NL} may also be expressed in the form

$$G_{NL} = \frac{g_m}{\beta}\frac{G_m}{g_m}, \tag{5.5–9}$$

where g_m/β is the small-signal input conductance looking into the base with the emitter grounded.

If the emitter, rather than the base, is driven as shown in Fig. 5.5–13, the equivalent linear loading on the input-tuned circuit is

$$G_{NL} = g_{in}\frac{G_m}{g_m}. \tag{5.5–10}$$

If either the base or emitter is driven and the bias is developed as shown in Fig. 5.4–1, the algebraic expressions for G_{NL} remain unchanged; however, $g_m = \alpha I_{EQ}q/kT$ and G_m/g_m must be obtained from Fig. 5.4–5 rather than Fig. 4.5–6.

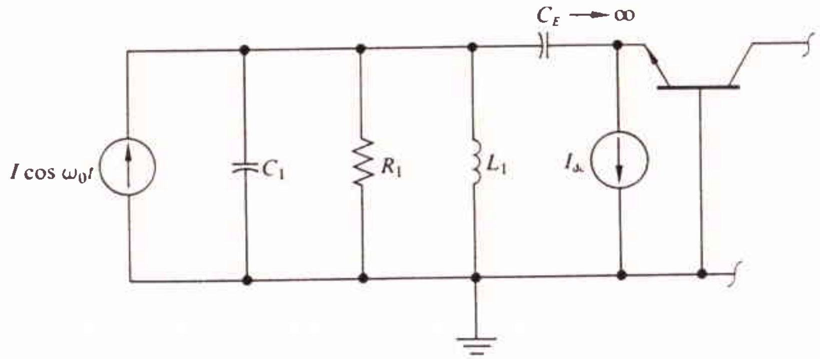


Figure 5.5–13

Example 5.5–5 For the circuit shown in Fig. 5.5–14, determine the output voltage $v_o(t)$ and the resistance seen by the input current in the steady state. Assume that $Q_T > 10$.

Solution. For the circuit of Fig. 5.5–14 we first note that, since Q_T is high and the circuit is driven by a current source at its resonant frequency, then $v_i(t)$ must have the form $v_i(t) = V_1\cos\omega_0 t$. We then observe that, since C is large, it appears as a battery V_{dc} in the steady state. Consequently the ideal diode-battery combination limits V_1 to V_{dc} (cf. Example 5.5–3) and in turn

$$v_o(t) = V_1(\cos\omega_0 t - 1).$$

Now, to evaluate V_1 as a function of I, we note that with $v_o(t)$ clamped to zero the diode conducts only at the peak of $v_o(t)$ and thus $i_D(t)$ flows in extremely narrow pulses; hence $i_D(t)$ may be represented as

$$i_D(t) \approx I_{Ddc}\left(1 + 2\sum_{n=1}^{\infty}\cos n\omega_0 t\right).$$

(For a periodic train of very narrow pulses, the amplitude of the harmonics is twice the average value.) However, since the average value of the capacitor current $i_o(t)$

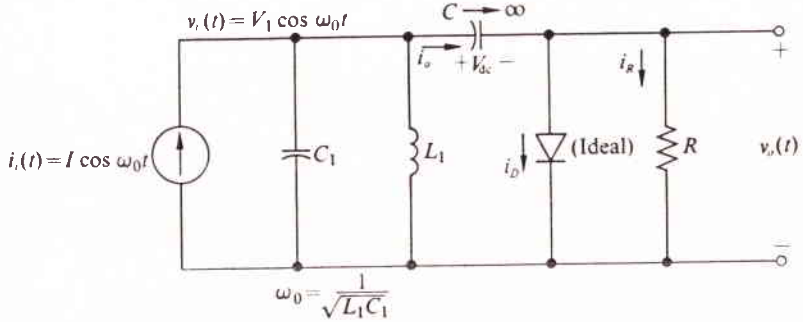

$v_i(t) = V_1\cos\omega_0 t$
$C \to \infty$
i_o
$+\ V_{dc}\ -$
i_R
+
$i_i(t) = I\cos\omega_0 t$
C_1
L_1
i_D
(Ideal)
R
$v_o(t)$
−
$\omega_0 = \frac{1}{\sqrt{L_1C_1}}$

Figure 5.5–14

ms of its MacLauren expansion, that is,

$$i_o(t) = a_0 + a_1 v_o(t) + a_2[v_o(t)]^2 + a_3[v_o(t)]^3 + \cdots . \tag{5.5–14}$$

the v_o-i_o characteristic is known analytically, then the coefficients a_n may be eval-ted in the form

$$a_n = \frac{1}{n!}\frac{di_o}{dv_o}\bigg|_{v_o=0} \tag{5.5–15}$$

owever, if the v_o-i_o characteristic is available only graphically, then a polynomial the desired order can be fitted to the characteristic by standard curve-fitting hniques.

Once a polynomial form for the v_o-i_o characteristic is obtained, it is a simple atter to determine the equivalent linear loading on a high-Q tuned circuit and the ultant level of $v_o(t)$. Specifically, if $v_o(t) = V_1 \cos \omega_0 t$ and we approximate $i_o(t)$ by olynomial of order 3, then

$$i_o(t) = a_0 + a_1 V_1 \cos \omega_0 t + a_2 V_1^2 \cos^2 \omega_0 t + a_3 V_1^3 \cos^3 \omega_0 t. \tag{5.5–16}$$

uation (5.5–16) may be arranged in Fourier-series form by use of the trigonometric ntities

$$\cos^3 \omega_0 t = \tfrac{3}{4}\cos \omega_0 t + \tfrac{1}{4}\cos 3\omega_0 t \qquad \text{and} \qquad \cos^2 \omega_0 t = \tfrac{1}{2} + \tfrac{1}{2}\cos 2\omega_0 t.$$

us

$$i_o(t) = a_0 + \frac{a_2 V_1}{2} + V_1\left(a_1 + \frac{3a_3}{4}V_1^2\right)\cos \omega_0 t$$

$$+ \frac{a_2}{2}V_1^2 \cos 2\omega_0 t + \frac{a_3}{4}V_1^3 \cos 3\omega_0 t, \tag{5.5–17}$$

m which we observe that

$$I_{o1} = V_1(a_1 + \tfrac{3}{4}a_3 V_1^2) \tag{5.5–18}$$

l

$$G_{NL} = \frac{I_{o1}}{V_1} = a_1 + \tfrac{3}{4}a_3 V_1^2. \tag{5.5–19}$$

Hence, if I_{i1} and G are known, then

$$I_{i1} = GV_1 + V_1(a_1 + \tfrac{3}{4}a_3 V_1^2) \tag{5.5–20}$$

l a cubic equation for obtaining V_1 is obtained which may be solved graphically analytically. For example, if

$$a_1 = G = 1\,\text{mmho}, \qquad a_3 = \frac{4\,\text{mmho}}{V^2}, \qquad \text{and} \qquad I_{i1} = 5\,\text{mA},$$

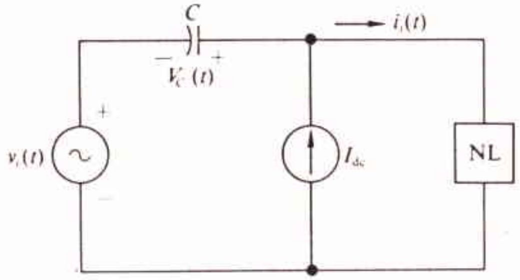
n the steady-state value of $v_o(t)$ is (1 V) cos $\omega_0 t$ and the equivalent linear loading he nonlinear load is $G_{NL} = 4$ mmho.

5.6 TRANSFER FUNCTION FOR LOW-INDEX AM INPUT

In this analysis we apply a low-index AM signal of the form

$$v_i(t) = [V_1 + v_1(t)] \cos \omega_0 t,$$

where $v_1(t) \ll V_1$, to the generalized capacitively coupled input circuit shown in Fig. 5.6–1 and determine the resultant AM variations on the envelope of the fundamental component of $i_i(t)$. We then adapt the analysis to the specific transistor amplifier shown in Fig. 5.3–1.



$$v_i(t) = V_1(t) \cos \omega t = [V_1 + v_1(t)] \cos \omega t$$
$$i_i(t) = I_0(t) + I_1(t) \cos \omega t + \cdots$$
$$I_0(t) = I_{dc} + i_0(t)$$
$$I_1(t) = I_1 + i_1(t)$$
$$V_C(t) = V_{dc} + v_0(t)$$

Fig. 5.6–1 Generalized capacitively coupled nonlinear input circuit.

Because of the nonlinear load, variations in the envelope of $v_i(t)$ produce variations not only in the envelope of $i_i(t)$ but also in V_C and the average value of $i_i(t)$. The form of these variations, which we assume to be small in comparison with the quiescent values of the respective parameters [i.e., the values obtained with $v_1(t) = 0$], is shown in Fig. 5.6–1. If small variations are assumed, it is apparent that

$$\begin{aligned} i_1(t) &= G_{11} v_1(t) + G_{10} v_0(t), \\ i_0(t) &= G_{01} v_1(t) + G_{00} v_0(t). \end{aligned} \tag{5.6–1}$$

and, in addition,

$$i_0(t) = -C\frac{dv_0(t)}{dt}, \tag{5.6–2}$$

where

$$G_{11} = \frac{\partial I_1}{\partial V_1}\bigg|_{Q\text{-pt}}, \qquad G_{10} = \frac{\partial I_1}{\partial V_C}\bigg|_{Q\text{-pt}},$$

$$G_{01} = \frac{\partial I_0}{\partial V_1}\bigg|_{Q\text{-pt}}, \qquad \text{and} \qquad G_{00} = \frac{\partial I_0}{\partial V_C}\bigg|_{Q\text{-pt}}$$

Q-pt denotes $V_C = V_{dc}$ and $V_1(t) = V_1$]. Thus, by expressing Eqs. (5.6–1) and (5.6–2) n Laplace transform notation and eliminating $i_0(t)$ and $v_0(t)$, we obtain

$$Y_1(p) = \frac{I_1(p)}{V_1(p)} = \frac{G_{11}\left[p + \dfrac{1}{C}\left(G_{00} - \dfrac{G_{10}G_{01}}{G_{11}}\right)\right]}{p + \dfrac{G_{00}}{C}}, \tag{5.6–3}$$

here $I_1(p) = \mathscr{L}[i_1(t)]$ and $V_1(p) = \mathscr{L}[v_1(t)]$. The transfer function $Y_1(p)$ permits the etermination of $i_1(t)$ once the form of $v_1(t)$ is known.

or the capacity coupled transistor amplifier of Fig. 5.3–1,

$$\begin{aligned} i_1(t) = i_E(t) &= I_{ES}e^{qV_C(t)/kT}e^{qV_1(t)/kT}\cos\omega t \\ &= I_{E_0}(t)\left\{1 + 2\frac{I_1[x(t)]}{I_0[x(t)]}\cos\omega t + \cdots\right\}, \end{aligned} \tag{5.6–4}$$

here

$$x(t) = \frac{qV_1(t)}{kT},$$

$$I_{E_0}(t) = I_{ES}I_0[x(t)]e^{qV_C(t)/kT},$$

nd

$$I_{E1}(t) = I_{ES}I_1[x(t)]e^{qV_C(t)/kT};$$

ence

$$G_{00} = \left.\frac{\partial I_{E0}}{\partial V_C}\right|_{Q\text{-pt}} = \frac{qI_{dc}}{kT} = g_{in}, \tag{5.6–5a}$$

$$\begin{aligned} G_{01} &= \left.\frac{\partial I_{E0}}{\partial V_1}\right|_{Q\text{-pt}} = \left.\frac{\partial I_{E0}}{\partial x}\cdot\frac{\partial x}{\partial V_1}\right|_{Q\text{-pt}} \\ &= g_{in}\frac{I_1(x)}{I_0(x)} \qquad \left[\text{Note:}\frac{dI_0(x)}{dx} = I_1(x)\right], \end{aligned} \tag{5.6–5b}$$

$$G_{10} = \left.\frac{\partial I_{E1}}{\partial V_C}\right|_{Q\text{-pt}} = g_{in}\frac{2I_1(x)}{I_0(x)}, \tag{5.6–5c}$$

nd

$$\begin{aligned} G_{11} &= \left.\frac{\partial I_{E1}}{\partial V_1}\right|_{Q\text{-pt}} = \left.\frac{\partial I_{E1}}{\partial x}\cdot\frac{\partial x}{\partial V_1}\right|_{Q\text{-pt}} \\ &= 2g_{in}\left[1 - \frac{I_1(x)}{xI_0(x)}\right] \qquad \text{(cf. Eq. A–11).} \end{aligned} \tag{5.6–5d}$$

Consequently, $Y_1(p)$ is given by

$$Y_1(p) = 2g_{in}\frac{\left[1 - \dfrac{I_1(x)}{xI_0(x)}\right]\left[\mathrm{p} + \dfrac{\lambda(x)}{\tau}\right]}{p + \dfrac{1}{\tau}}, \tag{5.6–6}$$

where $\tau = C/g_{in}$ and

$$\lambda(x) = 1 - \frac{\left[\dfrac{I_1(x)}{I_0(x)}\right]^2}{1 - \dfrac{I_1(x)}{xI_0(x)}}. \tag{5.6–7}$$

A plot of $\lambda(x)$ vs. x is given in Fig. 5.6–2. It can be shown (by differentiating Eq. (5.6–8)) that

$$\lambda(x) = \frac{\dfrac{d}{dx}\left[\dfrac{I_1(x)}{I_0(x)}\right]}{1 - \dfrac{I_1(x)}{xI_0(x)}}; \tag{5.6–8}$$

thus, for small values of p (or ω), $Y_1(p)$ is given by

$$Y_1(p) \approx Y_1(0) = g_{in}\frac{d}{dx}\frac{2I_1(x)}{I_0(x)}, \tag{5.6–9}$$

which is, of course, the admittance which relates small static variations in V_1 to I_{E1}.

From the pole-zero diagram for $Y_1(p)$ it is apparent that the maximum frequency component ω_m of $v_1(t)$ or $V_1(t)$ must be at least an octave below $\lambda(x)/\tau$, for any given value of x, to permit the use of the static transfer function relating I_{E1} and V_1 for the case of dynamic envelope variations. This restriction may be expressed in terms of the inequality

$$\frac{\omega_m\tau}{\lambda(x)} = \frac{\omega_m C}{g_{in}\lambda(x)} < 2. \tag{5.6–10}$$

Since $\lambda(x)$ decreases with increasing x, the maximum permissible value of the coupling capacitor which satisfies Eq. (5.6–10) also decreases. For example, if $x = 10$,

$$\lambda(x) = 0.0059 \qquad \text{and} \qquad \frac{\omega_m C}{g_{in}} < 0.0118.$$

However, from Eq. (5.3–20), $\omega C/g_{in} > 40\,\pi$; therefore, for correct operation of the circuit of Fig. 5.3–1, we require

$$\frac{\omega}{\omega_m} > 1.06 \times 10^4.$$

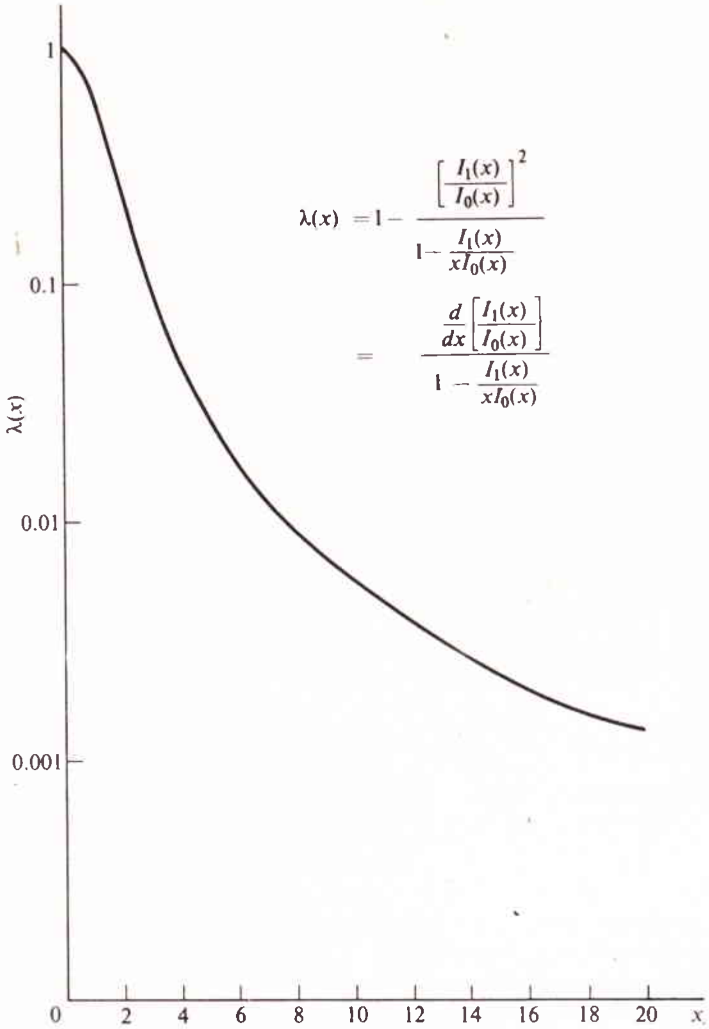


Fig. 5.6–2 Plot of $\lambda(x)$ vs. x.

ven though this restriction may be relaxed somewhat, we see that if the circuit is to unction as a narrowband limiter, which necessitates large values of x for its correct peration, there must exist a large separation between the maximum modulation requency and the carrier frequency. If this separation is not achieved, the circuit ils to limit properly, even though its static transfer characteristic indicates that it hould. Consequently, in limiter applications, the single-ended transistor is not articularly satisfactory and the differential pair, which may be operated without a oupling capacitor, finds more widespread use.

PROBLEMS

5.1 Show that a reasonable approximation for the "steady-state" values of the source current, $i_S(t)$, and the output voltage, $v_o(t)$, for the circuit of Fig. 5.P–1(a) is given by the results shown in Fig. 5.P–1(b). Assume that C is an ac short circuit.

5.2 Find the steady-state value of $v_o(t)$ in Fig. 5.P–2. How much average power is consumed in the 1 kΩ resistor after the capacitor voltage reaches equilibrium?

5.3 The repetitive waveshape shown is applied to the circuit in Fig. 5.P–3. Find the dc capacitor voltage and the shape and amplitude of the silicon diode current.

5.4 Sketch the collector current waveshape and find the collector voltage and the dc capacitor voltage in Fig. 4.P–4.

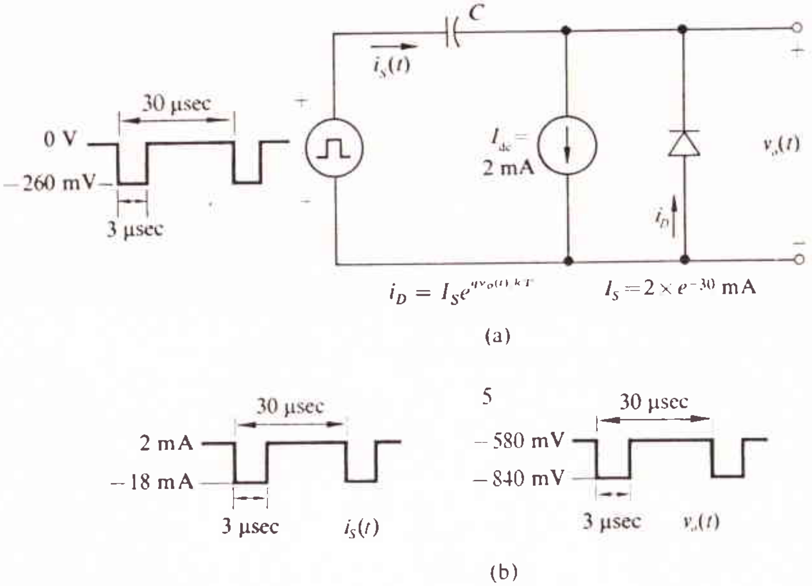


Figure 5.P–1

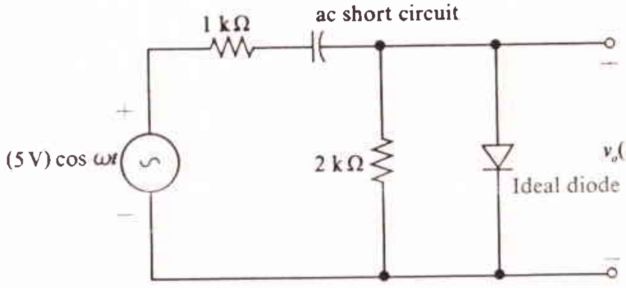


Figure 5.P–2

5.9 Explain how the results of Problems 5.7 and 5.8 would be modified if the radian frequency were reduced to 10^7 rad/sec.

5.10 Find the output voltage, $v_o(t)$, and the fundamental and third-harmonic diode currents for the circuit of Fig. 5.P–7.

5.11 Find the total harmonic distortion in $v_o(t)$ in the circuit of Fig. 5.4–7.

5.12 For the circuit shown in Fig. 5.P–8, determine an expression for $v(t)$ and $v_o(t)$.

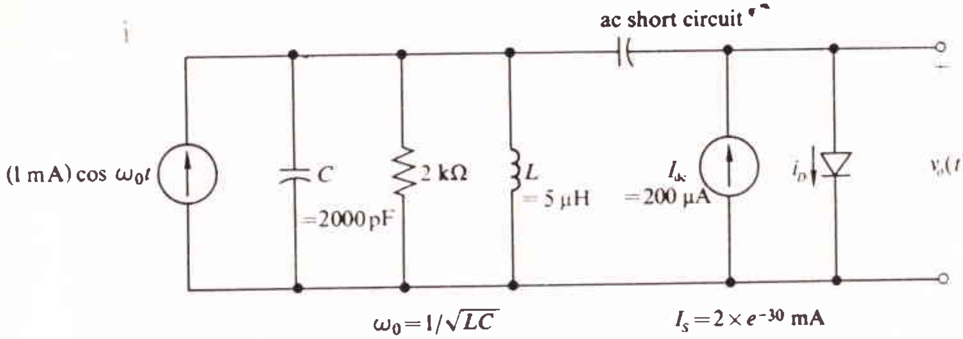


Figure 5.P–7

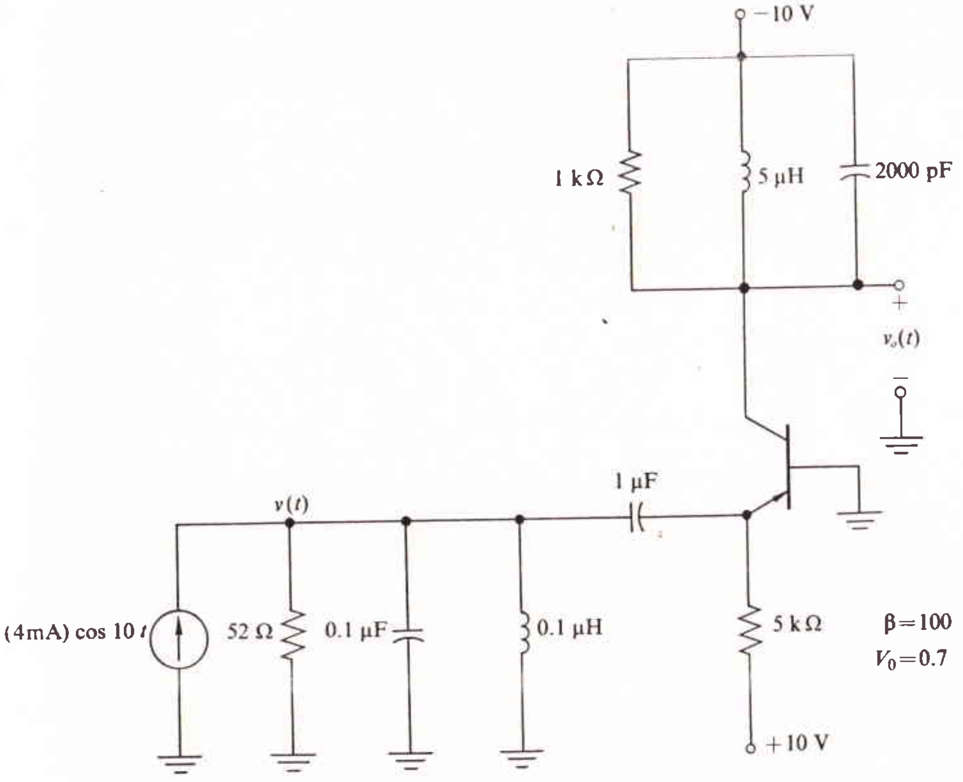


Figure 5.P–8

5.13 For the circuit shown in Fig. 5.P–9, find $v_1(t)$ and $v_o(t)$. Also sketch $i_{E2}(t)$.

5.14 For the circuit shown in Fig. 5.P–10, determine an expression for $v_o(t)$ and the equivalent linear loading across the tuned circuit for each of the following nonlinear loads and input current levels I_1. In each case find the THD present in $v_o(t)$:

a) $I_1 = 4$ mA, $i_D = I_S e^{qv_D/kT}$ (see Fig. 5.P–11a),

b) $I_1 = 2$ mA (see Fig 5.P–11b),

c) $I_1 = 2$ mA, $i_o = (20\,\mu V/V^3)v_o^3 + (10\,\mu V/V^5)v_o^5$,

d) $I_1 = 2$ mA (see Fig. 5.P–11c),

e) $I_1 = 1.5$ mA (see Fig. 5.P–11d).

5.15 Find the dc voltage across the input capacitor in the circuit of Fig. 4.P–11. Estimate the shape (conduction angle and peak value) of the collector current pulse for this case. Repeat for the case in which the source impedance is reduced to 5 Ω. What is the form and magnitude of $v_o(t)$ for this case? Does the transistor saturate?

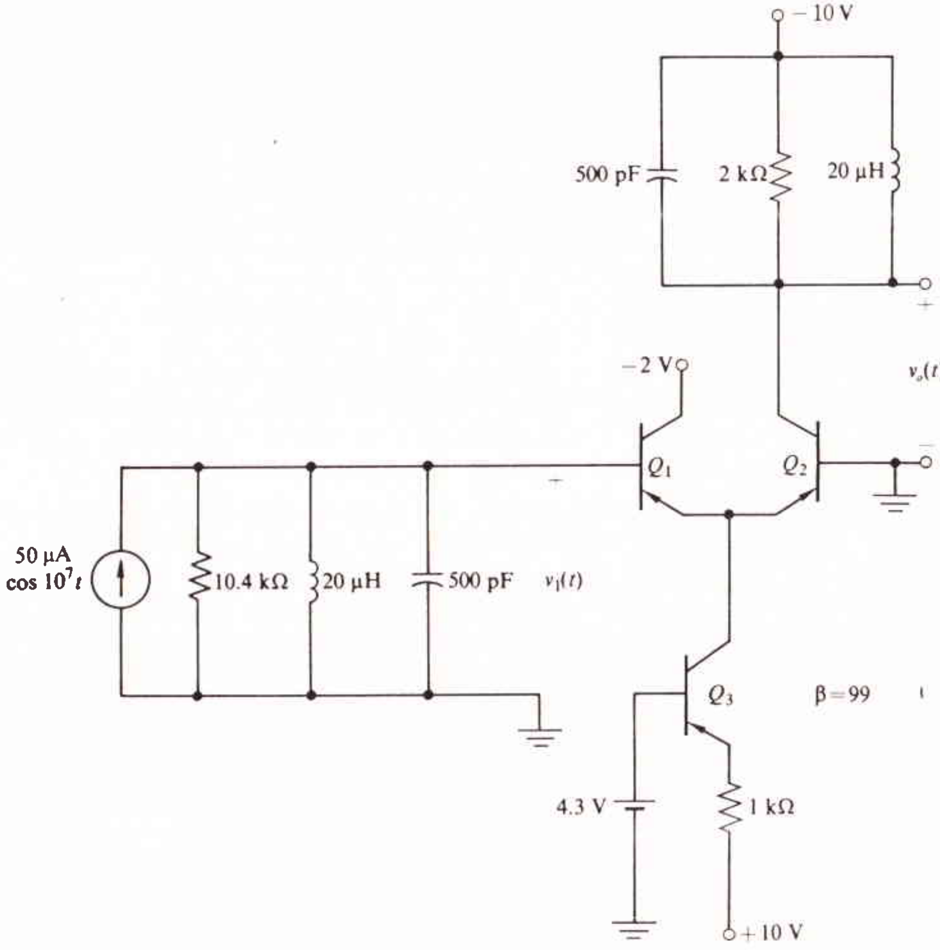


Figure 5.P–9

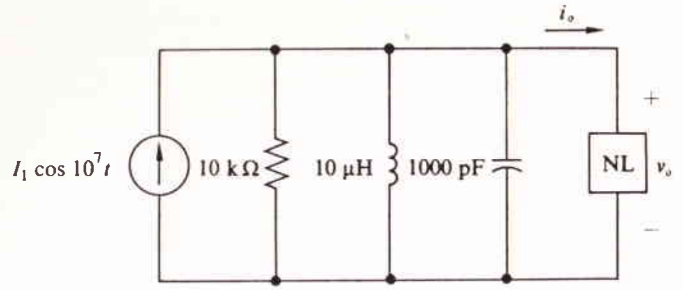


Figure 5.P–10

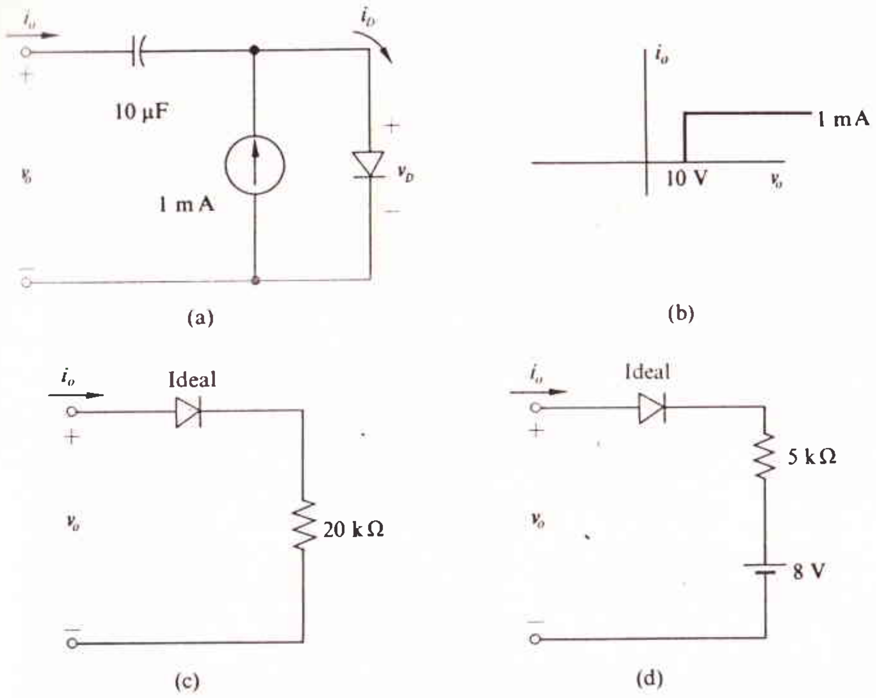


Figure 5.P–11

CHAPTER 6

SINUSOIDAL OSCILLATORS

In this chapter we combine the results of all the previous chapters to explore the design and analysis of a number of sinusoidal or almost sinusoidal oscillator circuits.

Once the waveshape of the output of such a circuit is specified as sinusoidal, then the two properties that remain to be defined are the operating frequency and the output amplitude. A good portion of this chapter will be devoted to examining the frequency- and amplitude-determining mechanisms of various circuits, the stabilities of these mechanisms with variables such as time, temperature, or supply voltage, and the interrelationships between amplitude and frequency.

There are methods of producing sinusoidal signals by filtering separately produced square waves or impulse chains or by the nonmemory shaping of triangular waves; however, the discussion of such circuits will be reserved for Chapter 11.

Both from the viewpoint of the techniques employed and from that of the practical division of circuits, it is convenient to consider most of our circuits as belonging to one of two classes. In the first class we combine a two-port active device with a two-port passive network in a feedback configuration. In the second type we combine a one-port active device in parallel with a one-port passive network. Obviously some circuits may be considered from either viewpoint. In spite of this, the division is still useful. We shall first consider the feedback type of circuitry. In subsequent sections we shall consider the one-port or "negative-resistance" circuits.

All sine-wave oscillators must, at the minimum, contain

a) an active device with power gain at the operating frequency (the circuit must supply not only an output signal but also its own input driving signal),
b) a frequency-determining element or network, and
c) an amplitude-limiting and stabilization mechanism.

In practical circuits these functions may become quite intermingled. However, initially we shall consider these operations separately so that the necessary or desirable properties of each one may become evident.

6.1 OPERATING FREQUENCY AND MINIMUM GAIN CONDITIONS FOR LINEAR-FEEDBACK OSCILLATORS

In order to sustain sinusoidal oscillations, a network must have a pair of complex conjugate poles in the right-half complex plane when power is applied at $t = 0$. These right-half plane poles, when excited by thermal noise or the step generated by closing the power switch, give rise to a sinusoidal output voltage with an exponentially growing envelope. If such a network has been designed to sustain a constant-amplitude sinusoidal output voltage, then, as the envelope of the noise-induced sine wave increases, it causes a change in the value of one or more of the network parameters (usually the amplification) in such a direction that the complex conjugate poles move toward the imaginary axis. Finally, at some predetermined amplitude of the growing sinusoid, the poles reach the imaginary axis and a constant-amplitude sinusoidal output is realized. If for some reason the output increases further, the poles move into the left-half plane and the output decreases toward the desired level, where again the poles lie on the imaginary axis.

We now see that the basic requirement for a sinusoidal oscillator consists of a pair of small-signal complex conjugate right-half plane poles, which determine the frequency of oscillation, plus a mechanism for moving these poles toward the imaginary axis as the envelope of the output waveform increases. This mechanism determines the steady-state oscillation amplitude. In order to obtain right-half plane poles, we require some form of feedback. Figure 6.1–1 illustrates a generalized feedback amplifier whose transfer function is given by

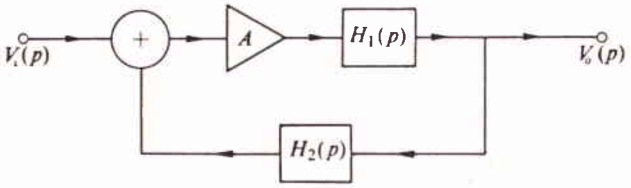


Fig. 6.1–1 Generalized feedback amplifier.

$$\frac{V_o(p)}{V_i(p)} = \frac{AH_1(p)}{1 - AH_1(p)H_2(p)} = \frac{AH_1(p)}{1 - A_L(p)}, \tag{6.1–1}$$

where $A_L(p) = AH_1(p)H_2(p)$ is defined as the "loop gain" of the amplifier. It is apparent that the poles of the feedback amplifier are the zeros or roots of

$$1 - A_L(p) = 0. \tag{6.1–2}$$

Thus if Eq. (6.1–2) has a pair of complex conjugate roots in the right-half plane, $v_o(t)$ will be a growing sinusoid even with $v_i(t) \equiv 0$.

The process of designing a sinusoidal oscillator now becomes clear. We select a suitable pole-zero pattern for

$$A_L(p) = AH_1(p)H_2(p)$$

which causes one pair of complex conjugate roots of Eq. (6.1–2) (and no other roots—otherwise unwanted oscillations might occur) to cross the imaginary axis at a predetermined frequency ω_0 as A (or $-A$) is increased. We also determine the minimum value $A_{\min}$ of A which places the roots on the imaginary axis, and choose A somewhat larger than this value to ensure self-starting. We then incorporate some nonlinear

mechanism which reduces A to $A_{\min}$ as the output oscillation grows toward a desired amplitude. Finally, we choose a network having the desired $A_L(p)$ and combine with it the desired nonlinear amplitude limiting.

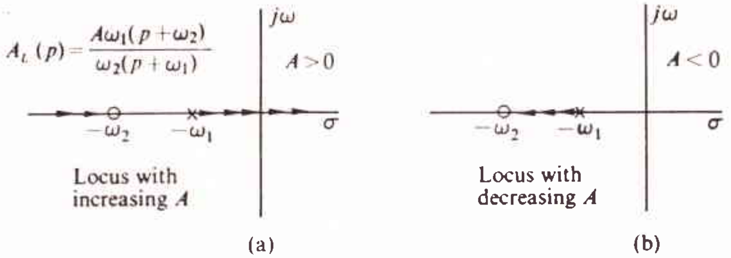
To arrive at pole-zero patterns for

$$A_L(p) = AH_1(p)H_2(p)$$

which are capable of producing oscillations, we start with the simplest pole-zero configurations and plot the loci of the roots of Eq. (6.1–2) as A is increased from zero to infinity and is decreased from zero to minus infinity. We then increase the complexity of the pole-zero diagram and repeat the root locus plot until a number of pole-zero diagrams capable of producing complex conjugate right-half plane roots are obtained. This approach immediately indicates that $A_L(p)$ must have two or more poles to cause Eq. (6.1–2) to have right-half plane complex conjugate roots. For example, with

$$A_L(p) = \frac{A\omega_1(p + \omega_2)}{\omega_2(p + \omega_1)}$$

the root locus is shown in Fig. 6.1–2 as A is increased (a) and then decreased (b) from zero. We see from the figure that a pole does indeed enter the right-half plane as A is increased. However, since it is a real-axis pole, the result is a growing exponential waveform and not a growing sinusoidal oscillation.


$A_L(p) = \frac{A\omega_1(p+\omega_2)}{\omega_2(p+\omega_1)}$
$j\omega$
$A>0$
$-\omega_2$
$-\omega_1$
σ
Locus with increasing A
$j\omega$
$A<0$
$-\omega_2$
$-\omega_1$
σ
Locus with decreasing A

(a) (b)

Fig. 6.1–2 Root locus of $A_L(p)$ with a single pole.

We also find that if $A_L(p)$ has two poles and no zeros, oscillations are impossible. The simplest pole-zero combination for $A_L(p)$ which is capable of producing right-half plane roots in Eq. (6.1–2) is the combination of two poles and one zero. The root locus for this case with

$$A_L(p) = \frac{A\omega_1 p}{(p + \omega_1)(p + \omega_2)}$$

is shown in Fig. 6.1–3, from which we observe that positive feedback, $A > 0$, is required for oscillations. It is apparent that for $A > 0$ the zero may move somewhat into the left-half plane without destroying the possibility of oscillation.

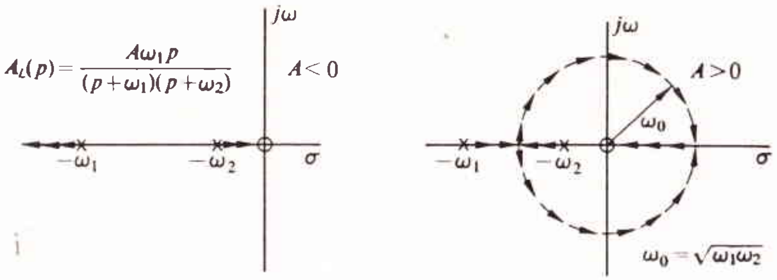


Fig. 6.1–3 Root locus of $A_L(p)$ with two poles and a zero.

To determine the frequency of oscillation ω_0 (i.e., the point at which the roots cross the imaginary axis) and the minimum gain A for oscillations (i.e., the value of A which just places the roots on the imaginary axis), we assume $A = A_{\min}$. With $A = A_{\min}$, $p = j\omega_0$ must be a solution of Eq. (6.1–2). Hence

$$A_L(j\omega_0) = 1, \tag{6.1–3}$$

or equivalently,

$$\mathscr{R}e\, A_L(j\omega_0) = 1, \tag{6.1–4a}$$

$$\mathscr{I}m\, A_L(j\omega_0) = 0. \tag{6.1–4b}$$

Equations (6.1–4a and b) provide a *necessary* condition for oscillation (this condition is known as the Barkhausen criterion) and thus a set of two equations for solving for $A_{\min}$ and ω_0. The Barkhausen criterion makes sense intuitively, since it states that at the frequency of oscillation ω_0 the signal must traverse the loop with no attenuation and with no phase shift.

With $A_L(p) = A\omega_1 p/(p + \omega_1)(p + \omega_2)$,

$$\mathscr{I}m\, A_L(j\omega_0) = \frac{A_{\min}\omega_1\omega_0(\omega_1\omega_2 - \omega_0^2)}{\omega_0^2(\omega_1 + \omega_2)^2 + (\omega_1\omega_2 - \omega_0^2)} = 0,$$

from which it is apparent that $\omega_0 = \sqrt{\omega_1\omega_2}$. In addition, Eq. (6.1–4a), with $\omega_0 = \sqrt{\omega_1\omega_2}$, yields

$$\mathscr{R}e\, A_L(j\omega_0) = \frac{A_{\min}\omega_1}{\omega_1 + \omega_2} = 1,$$

from which we obtain

$$A_{\min} = \frac{\omega_1 + \omega_2}{\omega_1}.$$

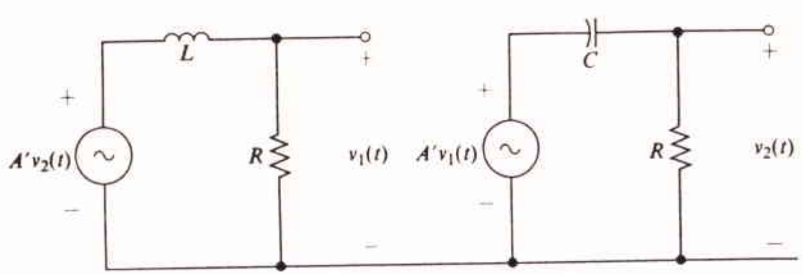


Fig. 6.1–4 Oscillator circuit with $A_L(p)$ having negative real poles and a zero at the origin.

Figure 6.1–4 illustrates a small-signal or linear model for a circuit with the pole-zero diagram shown in Fig. 6.1–3. For this circuit $A_L(p)$ is given by

$$A_L(p) = \frac{A'^2\left(\dfrac{R}{L}\right)p}{\left(\dfrac{R}{L} + p\right)\left(\dfrac{1}{RC} + p\right)};$$

thus $A = A'^2$, $\omega_1 = R/L$, and $\omega_2 = 1/RC$. If $1/RC = R/L$, then we have the special case where $\omega_1 = \omega_2 = \omega_0 = 1/\sqrt{LC}$ and $A_{\min} = 2$ or $A'_{\min} = \sqrt{2}$.

Another way to consider the circuit of Fig. 6.1–3 is to ask at what sinusoidal frequency the phase shift around the loop is zero. The first stage has a single pole at $-R/L$; hence its phase goes from zero to −90° as ω goes from zero to infinity. The second stage has a zero at the origin and a pole at $-1/RC$; hence its phase goes from +90° to zero as ω goes from zero to infinity. If $1/RC = R/L$, then the overall zero phase-shift point is at $\omega = 1/RC$. (At this point the magnitude of the phase shift from each circuit is 45°.) At this frequency each network has an attenuation of 3 dB; hence a gain per stage of $\sqrt{2}$ is required to raise the loop gain back to unity. This gain does not have to be spread through the circuit, but can be concentrated in one point.

Instead of two real-axis poles and a zero at the origin, $A_L(p)$ can be chosen to have a pair of complex conjugate left-half plane poles and a zero at or near the origin. The root locus for this case with

$$A_L(p) = \frac{Ap\omega_0'}{p^2 + 2\alpha p + \omega_0'^2} \tag{6.1–5}$$

is shown in Fig. 6.1–5. Here again we observe that oscillations are possible only for $A > 0$. However, this pole-zero configuration has two advantages over the configuration with real-axis poles. First, a smaller $A_{\min}$ is required to produce oscillations. Specifically, for this pole-zero pattern we find, employing the Barkhausen criterion, that

$$\omega_0 = \omega_0' \qquad \text{and} \qquad A_{\min} = \frac{2\alpha}{\omega_0'} = \frac{1}{Q_T},$$

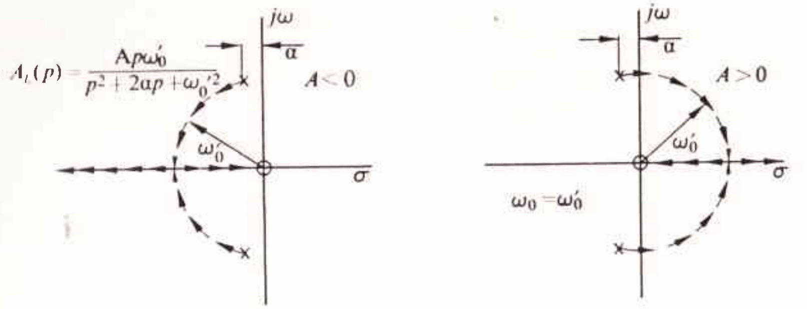


Fig. 6.1–5 Root locus for $A_L(p)$ with two complex conjugate poles and a zero at the origin.

vhere Q_T is the Q of the passive circuit within the feedback loop. Clearly, then, s Q_T increases, the required amplification of the active element within the feedback oop decreases.

A second advantage of the complex conjugate poles of $A_L(p)$ is increased fre- uency stability. If spurious poles and zeros should appear in the pole-zero pattern f $A_L(p)$ because of parasitic capacitance and inductance, the root locus is modified omewhat. However, if the poles of $A_L(p)$ are sufficiently close to the imaginary axis o start with, then the modified loci are still constrained to cross the imaginary axis elatively close to ω_0'.

This frequency stability may be observed from a different point of view. If the assive network has a high Q_T, its phase shift varies rapidly with frequency in the icinity of ω_0' (cf. Fig. 2.2–7). Therefore, spurious phase shifts introduced by parasitic lements within the feedback loop require only a small change in frequency from ω_0' o produce a compensating phase shift from the high-Q_T network and a net zero hase shift around the loop. Thus the higher the value of Q_T, the smaller the devia- ion from ω_0' to the frequency ω_0 at which zero loop phase shift exists and oscillations re possible.

For example, with $A_L(p)$ given by Eq. (6.1–5), the phase ϕ of $A_L(j\omega)$ in the vicinity f ω_0' is given by

$$\phi = -\tan^{-1}\frac{2Q_T(\omega - \omega_0')}{\omega_0'} \tag{6.1–6}$$

nd

$$\left.\frac{d\phi}{d\omega}\right|_{\omega_0'} = -\frac{2Q_T}{\omega_0'}.$$

Thus, with $Q = 10^4$ (typical of a quartz crystal), a 1° (0.0176-rad) shift in the overall oop phase shift leads to only an 87 rad/sec or a 13.8 Hz shift in frequency in an $\omega_0 = 10^8$ rad/sec oscillator.

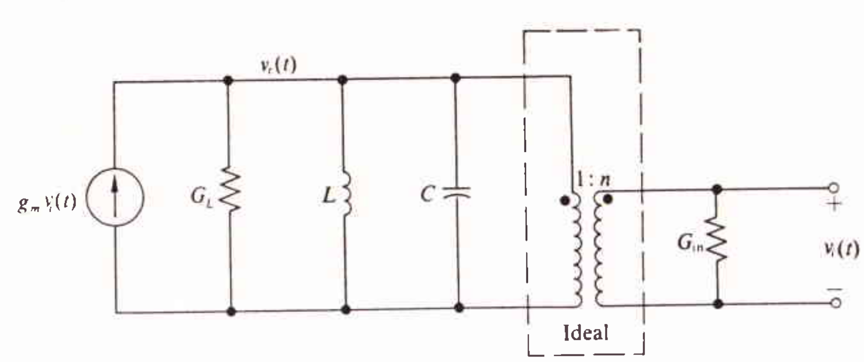


Fig. 6.1–6 Oscillator circuit with $A_L(p)$ having a pair of complex conjugate poles and a zero at the origin.

Figure 6.1–6 illustrates one of many small-signal oscillator circuits whose loop transfer $A_L(p)$ contains a pair of complex conjugate poles and a zero at the origin. For this circuit, $A_L(p)$ is given by

$$A_L(p) = \frac{(g_m n/G_T)(2\alpha p)}{p^2 + 2\alpha p + \omega_0'^2}, \tag{6.1–7}$$

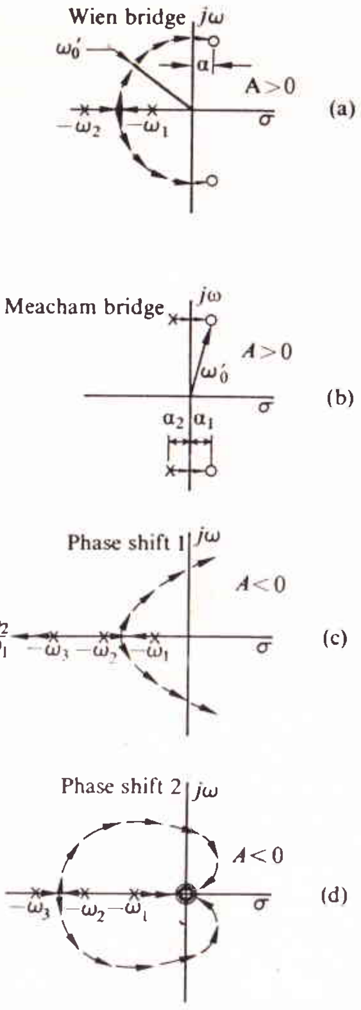
where $\omega_0' = 1/\sqrt{LC}$, $\alpha = G_T/C$, and $G_T = G_L + n^2 G_{in}$. This circuit could be realized by any of the transformerlike networks shown in Table 2.5–1 cascaded with a common base transistor, a common grid triode, or a common gate FET. Regardless of how the circuit of Fig. 6.1–6 is implemented, the minimum g_m required for oscillation at $\omega_0 = \omega_0'$ is

$$g_{m_{\min}} = \frac{G_T}{n}, \tag{6.1–8}$$

which decreases with decreasing G_T or, equivalently, with increasing $Q_T = \omega_0 C/G_T$.

Figure 6.1–7 illustrates the root loci for several other pole-zero patterns for $A_L(p)$ which are capable of oscillations. Included in the figure are the corresponding values of ω_0 and $A_{\min}$. The networks which yield loop gains with such pole-zero diagrams are explored in subsequent sections and in the problems at the end of the chapter. It should be noted in passing, however, that because of the closeness of the pole and zero in Fig. 6.1–7(b) (the Meacham bridge), spurious poles and zeros have very little effect on the j-axis crossing of the roots; thus this configuration should surpass the others in its ability to maintain a stable frequency.

In no case is the root locus plot necessary to solve an oscillator problem. It does help in understanding the action of the circuit and also in determining whether it is possible for a circuit to oscillate at a second undesired frequency. An alternative method of exploring the oscillating potential of a circuit is to plot the phase angle of $A_L(j\omega)$ vs. ω and to determine whether 0 or $\pm 2\pi$ of phase exists at a frequency where $|A_L(j\omega)|$ exceeds unity (cf. Eq. 6.1–3).



$$A_L(p)=\frac{A(p^2-2\alpha p+\omega_1\omega_2)}{(p+\omega_1)(p+\omega_2)}$$

$$A_{\min}=\frac{\omega_1+\omega_2}{2\alpha}$$

$$\omega_0=\omega_0'=\sqrt{\omega_1\omega_2}$$

$$A_L(p)=\frac{A(p^2-2\alpha_1 p+\omega_0'^2)}{(p^2+2\alpha_2 p+\omega_0'^2)}$$

$$A_{\min}=\frac{\alpha_2}{\alpha_1}$$

$$\omega_0=\omega_0'$$

$$A_L(p)=\frac{A\omega_1\omega_2\omega_3}{(p+\omega_1)(p+\omega_2)\,(p+\omega_3)}$$

$$A_{\min}=-(2+\frac{\omega_1}{\omega_3}+\frac{\omega_2}{\omega_3}+\frac{\omega_1}{\omega_2}+\frac{\omega_3}{\omega_2}+\frac{\omega_2}{\omega_1}+\frac{\omega_3}{\omega_1})$$

$$\omega_0=\sqrt{\omega_1\omega_2+\omega_2\omega_3+\omega_1\omega_3}$$

$$A_L(p)=\frac{Ap^3}{(p+\omega_1)(p+\omega_2)(p+\omega_3)}$$

$$A_{\min}=-(2+\frac{\omega_1}{\omega_3}+\frac{\omega_2}{\omega_3}+\frac{\omega_1}{\omega_2}+\frac{\omega_3}{\omega_2}+\frac{\omega_2}{\omega_1}+\frac{\omega_3}{\omega_1})$$

$$\omega_0=\sqrt{\frac{\omega_1\omega_2\omega_3}{\omega_1+\omega_2+\omega_3}}$$

Fig. 6.1–7 Root loci for several $A_L(p)$ capable of producing oscillations.

AMPLITUDE-LIMITING MECHANISMS

far we have ignored the problem of the determination of the amplitude at which oscillations stabilize. At least four different methods or combinations of methods amplitude stabilization are employed in practical circuits.

One or more temperature-sensitive resistors (a thermistor, a low-voltage light lb, or a semiconductor resistor) are used to build an attenuator. This power-sitive attenuator is constructed so that its attentuation λ increases with increasing ve level. It is inserted in the feedback loop in which the small-signal amplification

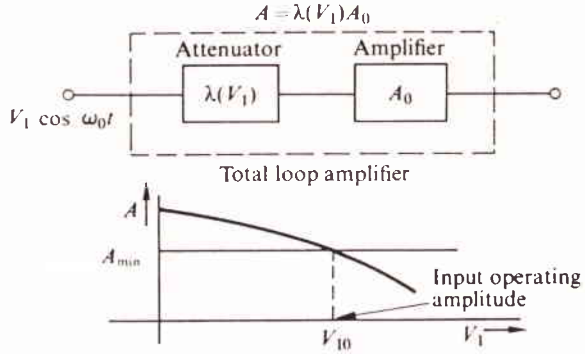


Fig. 6.2–1 "Nonlinear" attenuator as loop amplitude stabilizer.

exceeds the minimum required for oscillation. The result is a curve of loop gain scale factor A vs. attenuator input drive such as the one shown in Fig. 6.2–1.

The thermal time constant of the attenuator is made long in comparison with the period of the lowest frequency expected from the oscillator; hence in the steady state the attenuator is "linear." If, in addition, the signal levels in the active device are kept within the linear range of the device, low-distortion sinusoidal signals exist throughout the loop.

The long thermal time constant imposes a limitation both on the lowest frequency of operation and on the rate at which such oscillators can be swept or changed in frequency. This approach is widely used in audio- and video-frequency laboratory signal generators, particularly those of the RC network types. It is particularly applicable for adjusting R_1 and/or R_3 in the Wien bridge and Meacham bridge oscillators shown in later sections and in the problems at the end of the chapter.†

b) A second technique, which also produces low-distortion sinusoidal oscillations at all points throughout the feedback loop is the use of an amplitude detector (any of the types discussed in Section 10.2 or Section 10.4) to provide a control signal, proportional to the amplitude of the sinusoidal output, to inversely control the gain of some amplifier within the oscillator loop. Again the gain should decrease with increasing amplitude, so that a characteristic similar to Fig. 6.2–1 results. There is no necessity for either the amplitude detector or the gain-controlled amplifier to have a linear characteristic with respect to the envelope of the detector drive signal. The gain-controlled amplifier, however, must have a linear transfer characteristic if the oscillations throughout the loop are to have low distortion. As in case (a), the more steeply the variable-gain curve cuts through the required gain line, the better the amplitude stability will be.

---

† Hundreds of commercial thermistors (negative-temperature-coefficient resistors) exist. Resistance values for 25°C are available from 1 Ω or less up to 1 MΩ, with thermal time constants from 2 sec up to 200 sec or more. Not all combinations are available, nor are all units satisfactory from the reactive viewpoint for blind insertion into any network. Many units have only 20% tolerance and all are subject to variations with ambient temperature.

Any tungsten filament lamp with a rating of 6 W or less may be used as a positive-temperature-coefficient resistor. Hot resistances may be ten times as great as cold resistances. If the operating temperature is considerably above the ambient (but below incandescence), then the ambient variations will be much less significant than for a thermistor. Cold resistance (25°C) values range from below 1 Ω for 1 V flashlight bulbs to perhaps 400 Ω for a 3 W, 120 V light bulb.

Several of the references in Section 6.7 contain extensive data on such units. Later examples will contain illustrative data on some units.

c) A third and very common form of amplitude stabilization uses the nonlinearity of the active device to provide the amplitude-limiting feature for the circuit. In such cases the active device must be followed by a narrowband filter, which passes only the oscillation frequency, in order to remove the harmonics generated by the nonlinear active element and thus maintain a sinusoidal output signal. Most tuned oscillators similar to the one shown in Fig. 6.1–6 do provide the necessary filtering at the output of the active device.

If the nonlinear device is a common base transistor, then its effective transconductance at the oscillation frequency is $G_m(x)$, which is plotted vs. x in Fig. 4.5–6. Clearly, if for $x = 0$ ($V_1 = 0$) and $G_m(x) = g_{mQ}$ the small-signal value of $A_L(j\omega_0)$ is some number M which is greater than unity, then when power is applied to the oscillator at $t = 0$ the amplitude V_1 of the sinusoidal input to the transistor grows until it reaches a value V_{10} (or x_0) for which $A_L(j\omega_0)$ [or equivalently, $G_m(x)$] has decreased by a factor of M from its small-signal value. Therefore, x_0 and in turn $V_{10} = kTx_0/q$ may be obtained by determining at what point $G_m(x)/g_{mQ} = 1/M$ as illustrated in Fig. 6.2–2. For example, if $M = 2$, then from Fig. 4.5–6 $x_0 \doteq 3.3$.

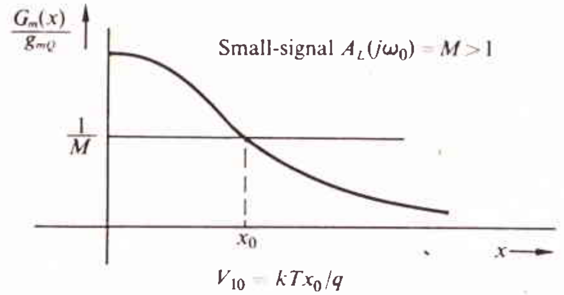


Fig. 6.2–2 Characteristic for self amplitude stabilization.

As we shall see, any of the nonlinear characteristics of fundamental current vs. sinusoidal drive voltage of Chapter 4 may be used as the basis for an amplitude stabilization scheme.

d) A fourth means of amplitude stabilization is presented by the problem of Example 5.5–3, where a diode-battery combination is used to limit the output sine wave from a tuned circuit.

While it is always possible to add an external diode and power supply, such additional circuitry may become available naturally in transistor or junction FET circuits when the collector-base or drain-gate diode is driven to the edge of conduction once per cycle. Figure 6.2–3 illustrates the situation.

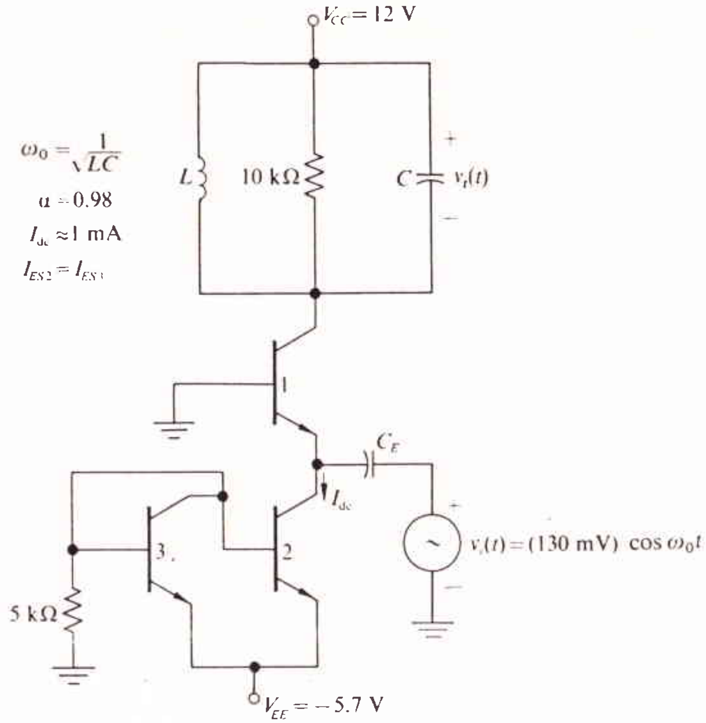


Fig. 6.2–3 Collector-base limiting.

From Eq. (4.5–13) and Table 4.5–1 we can evaluate the peak amplitude of the fundamental component of the collector current, I_1, in terms of I_{dc} and the peak amplitude of $v_i(t)$. In this case,

$$I_1 \approx 0.98 \times 1.79 = 1.75\text{ mA}.$$

Now if V_{CC} were large enough we would have a peak tank voltage of $R_T I_1 = 17.5$ V. Since $V_{CC} < 17.5$ V, the collector-base diode of the upper transistory conducts when $v_t(t)$ exceeds 12.7 V; this limits the output tank voltage to a sinusoidal peak amplitude of 12.7 V provided that $Q_T > 10$. Note that for the oscillator $v_i(t)$ is derived from $v_t(t)$; hence, once limiting from collector saturation occurs, we know the peak tuned-circuit voltage V_t, and in turn the feedback network yields the value of $v_i(t)$. With this value of $v_i(t)$ we can then determine I_1. The ratio V_t/I_1 yields the equivalent linear loading of the tuned-collector circuit, from which we can determine Q_T and, in turn, the distortion of the output sinusoid.

6.3 FREQUENCY STABILITY

One of the more important attributes of an oscillator is its ability to maintain an output frequency which is independent of changes in temperature, supply voltage, circuit loading, humidity, etc. Undesired frequency variations can, in general, be broken into two broad categories: "direct" frequency variations and "indirect" frequency variations. Direct frequency variations result from changes in the parameters which directly control ω_0. For example, for the oscillator of Fig. 6.1–6 ω_0 is given by $\omega_0 = 1/\sqrt{LC}$; hence, if L or C varies with temperature, ω_0 also varies with temperature. Quantitatively this variation in ω_0 may be expressed in the form

$$\begin{aligned}\frac{\Delta\omega}{\omega_0} &= \frac{1}{\omega_0}\left(\frac{\partial\omega_0}{\partial L}\Delta L + \frac{\partial\omega_0}{\partial C}\Delta C\right) \qquad (6.3\text{–}1)\\ &= -\frac{1}{2}\left(\frac{\Delta L}{L} + \frac{\Delta C}{C}\right),\end{aligned}$$

where $\Delta\omega$ is the variation of the oscillation frequency resulting from variations of ΔL and ΔC in the inductance and capacitance respectively. Clearly a 1% increase in L or C results in a $\frac{1}{2}$% decrease in ω_0. Therefore, for the case of an LC oscillator of the form of Fig. 6.1–6, the main frequency-determining elements, the inductance and the capacitance, must be stabilized against temperature variations, stray loading, etc.

Temperature stability is usually accomplished by trying to choose an inductor and capacitor (or resistors and capacitors in an RC oscillator) with equal and opposite temperature coefficients, thereby causing the variations to cancel in Eq. (6.3–1).† Stray capacitive or inductive loading is usually minimized by coupling the LC tuned circuit to the active element by means of a step-down transformer, as shown in Fig. 6.4–9. Still further frequency stability against temperature may be accomplished by placing the main frequency-determining elements in an oven. This is often done when a crystal is employed in lieu of the LC tuned circuit in the oscillator.

In general, direct frequency variation of ω_0 may be expressed as

$$\frac{\Delta\omega}{\omega_0} = \frac{1}{\omega_0}\sum_{i=1}^{n}\frac{\partial\omega_0}{\partial\alpha_i}\Delta\alpha_i, \qquad (6.3\text{–}2)$$

where the α_i's are the various circuit parameters on which ω_0 depends. Clearly the highest direct frequency stability is achieved by choosing an oscillator configuration for which ω_0 depends on the fewest number of parameters (α_i) and then stabilizing these parameters against environmental changes. Alternatively the parameters may be chosen in such a way that their net variation with ambient changes results in a small value $\Delta\omega$.

---

† Typical L and C temperature coefficients are tabulated by Eisenberg: B. R. Eisenberg, "Frequency Stability of a Clapp VCO," *IEEE Transactions on Instrumentation and Measurement*, **IM–18**, No. 3, pp. 221 ff. (Sept. 1969).

The oscillator whose pole-zero diagram is shown in Fig. 6.1–5 and the Meacham bridge oscillator (Fig. 6.1–7b) are two examples of oscillators with values of ω_0 which depend on a small number of parameters. Consequently these pole-zero diagrams represent optimum configurations for achieving a high direct frequency stability.

Indirect frequency variations are a result of parasitic reactances which introduce additional poles and zeros, with unknown locations, into the pole-zero diagram of $A_L(p)$, thereby changing the frequency of oscillation. Unfortunately many of these reactances are functions of temperature and supply voltage (e.g., output capacitance of a transistor is a function of the collector-to-emitter voltage and therefore the supply voltage).

To see how these parasitic poles and zeros affect ω_0, we recall that ω_0 is the radian frequency at which the net phase shift of $A_L(p)$ is zero. Since the parasitic poles and zeros introduce additional phase shift (let us say $\Delta\phi$), the frequency of oscillation must shift by an amount $\Delta\omega$ which causes the main frequency-determining elements of $A_L(p)$ (the known poles and zeros) to supply a phase shift of $-\Delta\phi$ and thus maintain zero phase shift around the feedback loop.

For example, if we have a crystal oscillator that operates near the nominal series resonant frequency, ω_0, of the crystal, then if the rest of the circuit suddenly introduces an additional phase shift of +1 deg, the operating frequency shifts until the transmission through the crystal produces an additional compensating −1 deg of phase shift. It is apparent that if the crystal has a variation of phase angle with frequency of 1 deg/Hz, then only a 1 Hz shift in frequency results.

Thus the greater the phase variation $\Delta\phi$ of $A_L(j\omega)$ for a change in frequency $\Delta\omega$ from ω_0, the greater the indirect frequency stability. Consequently we define the indirect frequency stability S_F as

$$S_F = \frac{\Delta\phi}{\Delta\omega/\omega_0} \approx \omega_0\frac{d\phi}{d\omega}\bigg|_{\omega=\omega_0}, \qquad (6.3\text{–}3)$$

where $\Delta\phi$ is the phase variation of $A_L(j\omega)$ due to a change in frequency $\Delta\omega$ from ω_0. The larger the value of S_F, the better the circuit functions in combating parasitic changes in frequency. Specifically, if parasitic elements in the feedback loop change the phase by $\Delta\phi$, the resultant change in frequency from ω_0 is given by

$$\Delta\omega = -\frac{\omega_0\,\Delta\phi}{S_F}. \qquad (6.3\text{–}4)$$

Note, however, that S_F predicts variations in ω_0 only for those cases where the main frequency-determining poles and zeros of $A_L(p)$ are stable; i.e., direct frequency stability has been accomplished.

In order to compare circuits, it is instructive to evaluate S_F for the several oscillators for which the pole-zero diagrams of $A_L(p)$ were presented in Section 6.1. In each case we must evaluate

$$\omega_0\frac{d\phi}{d\omega}\bigg|_{\omega_0};$$

wever, since

$$\phi = \sum_{i=1}^{n} \phi_i,$$

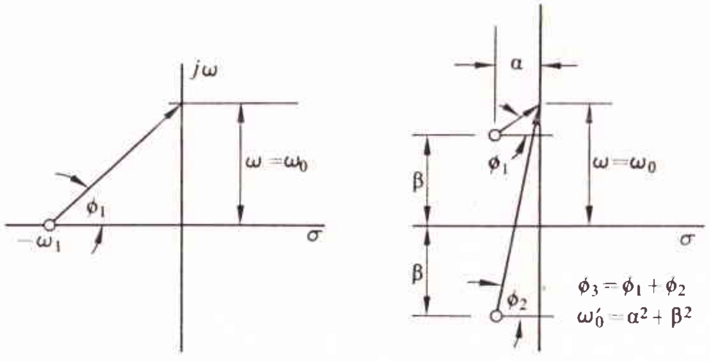
ere ϕ_i's are the angles contributed by the individual poles and zeros of $A_L(p)$, en

$$S_F = \omega_0 \sum_{i=1}^{n} \left.\frac{d\phi_i}{d\omega}\right|_{\omega_0}. \tag{6.3–5}$$

nsequently, by evaluating

$$S_{Fi} = \omega_0 \left.\frac{d\phi_i}{d\omega}\right|_{\omega_0}$$

a single real-axis pole and zero, a pair of complex conjugate poles, and a pair of nplex conjugate zeros, we may apply superposition to obtain S_F for any of the le-zero diagrams for $A_L(p)$.



(a) (b)

. 6.3–1 Pole-zero diagrams of a single real-axis zero and a pair of complex conjugate zeros.

The phase angle for the real-axis zero at $-\omega_1$ shown in Fig. 6.3–1(a) is given by

$$\phi_1 = \tan^{-1}\frac{\omega}{\omega_1}; \tag{6.3–6}$$

s

$$S_{F1} = \omega_0 \left.\frac{d\phi_1}{d\omega}\right|_{\omega_0} = \frac{\omega_0/\omega_1}{1 + (\omega_0/\omega_1)^2}. \tag{6.3–7}$$

eal-axis pole $-\omega_1$ would have the same magnitude of S_{F1} but a negative sign. s apparent that, if the real-axis pole or zero lies at the origin ($\omega_1 = 0$), its net ntribution to the indirect frequency stability is zero. This is so because the phase

---

of the pole or zero remains constant at $+\pi/2$ or $-\pi/2$ respectively, independent of the variation in ω. In addition, it is apparent that if ω_1 is close to infinity its contribution to S_F is small.

The phase angle ϕ_3 of the pair of complex conjugate zeros shown in Fig. 6.3–1(b) is given by

$$\phi_3 = \phi_1 + \phi_2 = \tan^{-1}\frac{\omega - \beta}{\alpha} + \tan^{-1}\frac{\omega + \beta}{\alpha}; \tag{6.3–8}$$

thus

$$S_{F3} = \omega_0 \left.\frac{d\phi_3}{d\omega}\right|_{\omega_0} = \frac{2\alpha\omega_0(\omega_0'^2 + \omega_0^2)}{(\omega_0'^2 - \omega_0^2)^2 + 4\alpha^2\omega_0^2}, \tag{6.3–9}$$

where $\omega_0'^2 = \alpha^2 + \beta^2$. For the usual case where $\omega_0' = \omega_0$ (the Wien bridge and the Meacham bridge), S_{F3} simplifies to

$$S_{F3} = \frac{\omega_0'}{\alpha} = 2Q_{Tz}, \tag{6.3–10}$$

where $Q_{Tz} = \omega_0'/2\alpha$ is the effective Q of the zeros. Again, for a pair of complex conjugate poles, the magnitude of S_{F3} is unchanged but the sign is reversed. Similarly the sign would be reversed if the zeros were in the right-half complex plane with the same Q_{Tz}.

If we apply the above results to the oscillator for which $A_L(p)$ consists of two negative real-axis poles at $-\omega_1$ and $-\omega_2$ and a zero at the origin as shown in Fig. 6.1–3, we obtain

$$S_F = -\frac{\omega_0/\omega_1}{1 + (\omega_0/\omega_1)^2} - \frac{\omega_0/\omega_2}{1 + (\omega_0/\omega_2)^2} + 0. \tag{6.3–11}$$

If we introduce the fact that $\omega_0^2 = \omega_1\omega_2$, then S_F simplifies to

$$S_F = -\frac{2\sqrt{\omega_1\omega_2}}{\omega_1 + \omega_2} = -\frac{2\sqrt{\omega_1/\omega_2}}{1 + \omega_1/\omega_2}. \tag{6.3–12}$$

The minus sign merely implies that ϕ is decreasing with increasing ω; hence a positive $\Delta\phi$ from the rest of the circuit requires a negative $\Delta\phi$ from the frequency-determining elements and a positive shift in frequency. A plot of $-S_F$ vs. ω_1/ω_2 is given in Fig. 6.3–2; from this we observe that S_F reaches its maximum value of -1 when the two poles of $A_L(p)$ lie at the same point. In addition, the maximum is sufficiently broad so that even if the pole positions differ by a factor of 14, S_F has decreased only by a factor of 2 from its maximum value.

For the oscillator for which $A_L(p)$ has a pair of complex conjugate poles at $-\alpha \pm j\beta$ (where $\alpha^2 + \beta^2 = \omega_0'^2 = \omega_0^2$) and a zero at the origin, S_F may be written directly from Eq. (6.3–10) as

$$S_F = -2Q_T, \tag{6.3–13}$$

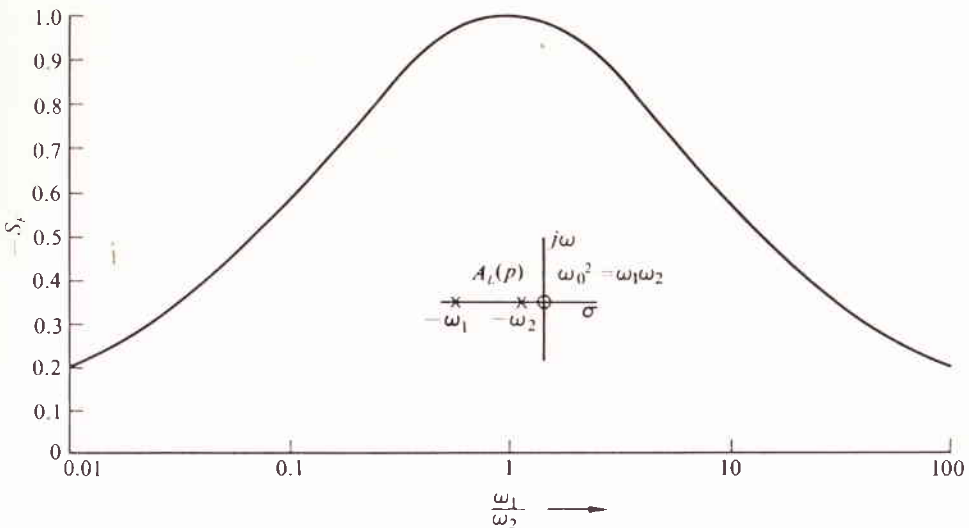


Fig. 6.3–2 Plot of $-S_F$ vs. ω_1/ω_2 for $A_L(p) = A\omega_1 p/(p + \omega_1)(p + \omega_2)$.

where Q_T is the Q of the passive network yielding $A_L(p)$. Clearly, as anticipated in the previous section, the higher the value of Q_T, the greater the immunity to changes in frequency due to parasitic elements.

For the Wien bridge oscillator for which $A_L(p)$ is shown in Fig. 6.1–7(a) and for which $\omega_0^2 = \omega_0'^2 = \omega_1\omega_2$, S_F may be written directly as the superposition of Eqs. (6.3–10) (with a negative sign) and (6.3–12); thus

$$S_F = -\frac{2\sqrt{\omega_1\omega_2}}{\omega_1 + \omega_2} - 2Q_{Tz}. \qquad (6.3\text{–}14)$$

It is quite apparent that the indirect frequency stability of the Wien bridge oscillator exceeds that of the oscillator with two negative real-axis poles by $2Q_{Tz}$, while its frequency stability is essentially the same as that of the parallel RLC oscillator with comparable Q_T.

For the Meacham bridge oscillator for which $A_L(p)$ is shown in Fig. 6.1–7(b) and for which $\omega_0 = \omega_0'$, S_F may be immediately written as

$$S_F = -2(Q_T + Q_{Tz}), \qquad (6.3\text{–}15)$$

where Q_T and Q_{Tz} are the Q's of the poles and zeros respectively. As we shall see, in many cases

$$Q_{Tz} \approx \frac{Q_T A}{4},$$

where A is the voltage amplification of the active device within the oscillator loop. Hence with $A = 100$ we can obtain an S_F 26 times that of the pole pair alone. Thus,

if the poles of $A_L(p)$ are the result of an RLC network with $Q_T = 100$, then $S_F \approx -5200$. If the poles are due to a crystal with $Q_T = 5000$, S_F can be increased to $S_F \approx 260{,}000$.

For either of the phase shift oscillators for which $A_L(p)$ is given in Fig. 6.1–7 (c and d), S_F may be immediately written as

$$S_F = -\sum_{i=1}^{3} \frac{\omega_0/\omega_i}{1 + (\omega_0/\omega_i)^2}, \qquad (6.3\text{–}16)$$

which, for the special case† where

$$\omega_1 = \omega_2 = \omega_3 = \frac{\omega_0}{\sqrt{3}} \quad \text{or} \quad \omega_1 = \omega_2 = \omega_3 = \sqrt{3}\omega_0$$

[Figs. 6.1–7(c) and (d) respectively], reduces to

$$S_F = -\frac{3\sqrt{3}}{4} = -1.3.$$

This value is significantly less than that obtainable with the Wien bridge or Meacham bridge oscillator or with the oscillator shown in Fig. 6.1–6 with a high value of Q_T. In addition, for the phase shift oscillators ω_0 depends on ω_1, ω_2, and ω_3; hence these oscillators have poor direct frequency stability.

Internal phase shifts are caused not only by reactances within active devices or amplifiers but by dc blocking capacitors, by transformer couplings, and by the presence of harmonic terms in the active device voltage or current ouput. In all cases, the higher the S_F, the less troublesome these effects will be.

As we proceed through this chapter it will be useful to bear in mind that it is generally desirable in any oscillator circuit to provide the maximum possible isolation between the device and the network. For one thing, this isolation will increase the circuit's frequency stability, since the frequency-affecting device reactances are normally functions both of the Q-point and of temperature. Most useful oscillator circuits achieve this isolation by what amounts to a gross impedance mismatching between the device input and the network output impedance and between the network input impedance and the device output impedance.

From this point of view oscillator circuits can be classified into the four forms shown in Fig. 6.3–3. In this figure the device is shown as driving the network, which in turn works into the input impedance of the device. The term "current-driven" implies that the network input impedance is small in comparison to the output impedance of the device; the term "voltage-driven" implies that the network input impedance is large in comparison to the output impedance of the device.

All such classifications are relative. Thus a 50 Ω device input impedance may be considered as an open circuit if the network has an output impedance of 5 Ω.

Most of the circuits of the next three sections will be of the type shown in Fig. 6.3–3(a). The bridge circuits of Section 6.8 are normally of the form shown in Fig. 6.3–3(c), while certain of the crystal oscillators, as well as some of the circuits of Section 6.10, illustrate the other two forms.

---

† This case results in the maximum value of S_F.

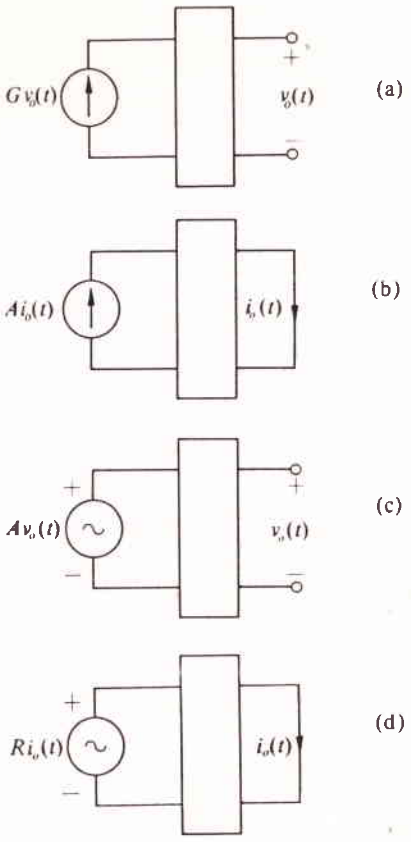
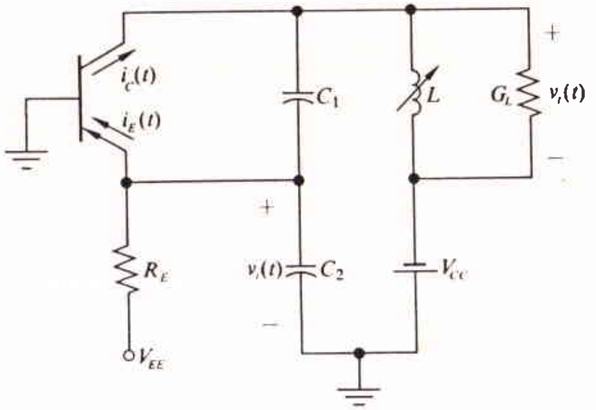


Fig. 6.3–3 (a) Current-driven network working into an "open" circuit. (b) Current-driven network working into a "short" circuit. (c) Voltage-driven network working into an "open" circuit. (d) Voltage-driven network working into a "short" circuit.

This general classification is often useful when one is considering the use of a ew device in an oscillator circuit or when one is considering a possible new circuit onfiguration. In such cases it is usually desirable to go back to the basic principles ather than merely to try to "convert" some existing circuit.

4 SELF-LIMITING SINGLE-TRANSISTOR OSCILLATOR

he majority of actual oscillators in common application are self-limiting singleansistor oscillators of the form shown in Fig. 6.4–1. In this section we shall examine ne properties of this type of oscillator, and in subsequent sections we shall extend ur analysis to the cases where the active element is replaced by a differential pair or a nction field effect transistor.



UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION
Y BIBLIOTECA
MONTEVIDEO - URUGUAY

Fig. 6.4–1 Prototype single-transistor Colpitts oscillator.

Although the frequency-determining network in this (Colpitts) oscillator is a tapped capacitive transformer, any of the transformerlike networks shown in Table 2.5–1 can be employed without altering the analysis. Some of the other configurations are shown in Fig. 6.4–2.

For the circuit of Fig. 6.4–1 the quiescent emitter current is given by

$$I_{EQ} = \frac{V_{EE} - V_{BEQ}}{R_E} = \frac{V_\lambda}{R_E},$$

where $V_{BEQ} \approx 0.65$ V for a silicon transistor and 0.22 V for a germanium transistor. Thus the small-signal input conductance at the emitter is given by

$$g_{\text{in}Q} = \frac{qI_{EQ}}{kT},$$

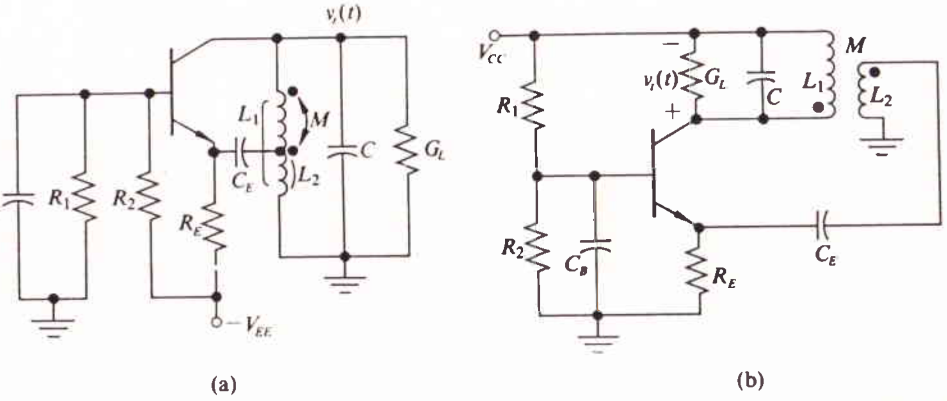
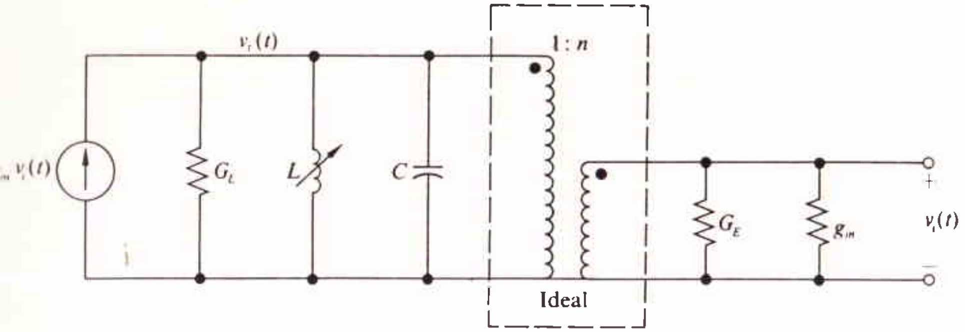


Fig. 6.4–2 Other configurations for single-transistor oscillators, (a) Hartley oscillator, (b) Tuned-collector oscillator.



$nQ_T{\cdot}Q_E > 100$ $\quad C = C_1C_2/(C_1 + C_2)$ $\quad n = C_1/(C_1 + C_2)$
$Q_E > 10$ $\quad g_{in} = g_m/\alpha$ $\quad \omega_0 = 1/\sqrt{LC}$

Fig. 6.4–3 Small-signal model for single-transistor oscillator.

nd the small-signal transconductance has the form

$$g_{mQ} = \alpha g_{inQ}.$$

f we assume $Q_T > 10$, $Q_E > 10$, and $nQ_T{\cdot}Q_E > 100$, then the capacitive transformer nay be replaced by the transformer model of Table 2.5–1 to yield the small-signal scillator circuit shown in Fig. 6.4–3. Clearly the small-signal model is identical to he model shown in Fig. 6.1–6, from which we observe that growing oscillations ccur with a frequency of $\omega_0 = 1/\sqrt{LC}$ provided that

$$g_{mQ} > g_{mQ_{\min}} = \frac{G_L + n^2(G_E + g_{mQ_{\min}}/\alpha)}{n}, \tag{6.4–1}$$

r equivalently,

$$g_{mQ} > g_{mQ_{\min}} = \frac{G_L + n^2G_E}{n(1 - n/\alpha)}. \tag{6.4–2}$$

the inequality of Eq. (6.4–2) is satisfied, the oscillation grows until the transistor onlinearities reduce $A_L(j\omega_0)$ to unity, at which level the oscillation stabilizes.

To determine at what amplitude (V_t) the ac collector voltage $v_t(t)$ stabilizes, we ssume

) that collector saturation does not occur and

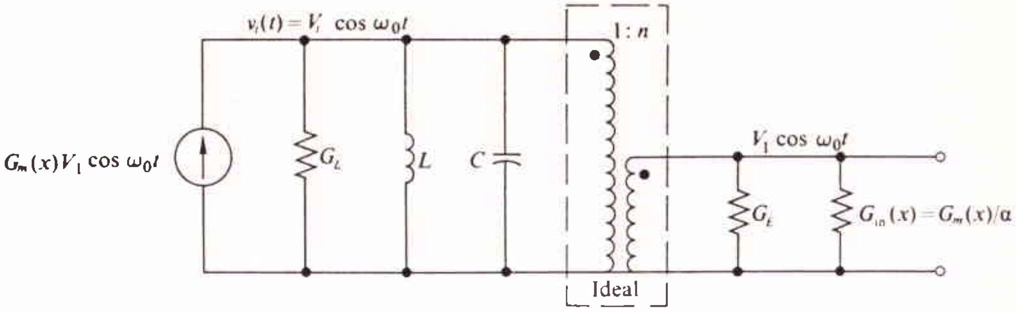
) that the passive circuit Q_T is sufficiently high so that for large-signal operation the second and higher harmonics of $v_i(t)$ and $v_t(t)$ are negligible, and consequently the capacitive transformer can be replaced by its ideal transformer model given in Table 2.5–1.

Vith these assumptions, on a large-signal basis,

$$v_t(t) = V_t \cos \omega_0 t$$

nd

$$v_i(t) = V_1 \cos \omega_0 t + V_{dc}, \tag{6.4–3}$$


$v_i(t) = V_i \cos \omega_0 t$
$1:n$
$G_m(x)V_1 \cos \omega_0 t$
G_L
L
C
$V_1 \cos \omega_0 t$
G_E
$G_{in}(x) = G_m(x)/\alpha$
Ideal

$C = C_1C_2/(C_1 + C_2)$, $\quad n = C_1/(C_1 + C_2)$, $\quad \omega_0 = 1/\sqrt{LC}$

Fig. 6.4–4 Large-signal model for the circuit of Fig. 6.4–1.

where $V_1 = nV_t$ and V_{dc} is the average value of the emitter-to-base voltage. In addition, the circuit for obtaining $v_t(t)$ may be modeled as shown in Fig. 6.4–4. In this model the nonlinear transistor input characteristic has been replaced by its equivalent linear conductance $G_m(x)/\alpha$(cf. Example 5.5–4) and only the fundamental component of collector current $G_m(x)V_1 \cos \omega_0 t$ has been retained, since all other harmonic components are effectively shunted to ground by the high-Q_T tuned circuit. The transconductance $G_m(x)$ is given by Eq. (5.4–17).

It is apparent that the large-signal model of Fig. 6.4–4 is identical to the small-signal model of Fig. 6.4–3 except that $G_m(x)$ replaces g_{mQ}; therefore, $A_L(j\omega_0) = 1$ for that value of $x = qV_1/kT$ for which

$$G_m(x) = g_{mQ_{\min}} = \frac{G_L + n^2G_E}{n(1 - n/\alpha)}, \tag{6.4–4}$$

or equivalently,

$$\frac{G_m(x)}{g_{mQ}} = \frac{2I_1(x)}{xI_0(x)}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right] = \frac{G_L + n^2G_E}{g_{mQ}n(1 - n/\alpha)}. \tag{6.4–5}$$

By plotting the constant

$$\frac{G_L + n^2G_E}{g_{mQ}n(1 - n/\alpha)}$$

on the characteristic of Fig. 5.4–5, we obtain (with the appropriate value of V_λ) the value of x, and in turn V_1, for which the oscillation stabilizes. With V_1 and n known, $V_t = V_1/n$ is readily determined and the problem is solved.

In almost all self-limiting oscillators, Q_T is high (to keep the output voltage sinusoidal) and $n \ll 1$ (since the output voltage amplitude is usually required to be large in comparison with V_1). With these two conditions ($Q_T \gg 1$ and $n \ll 1$), certain simplifications result. The first is that Eq. (6.4–5) may be closely approximated as

$$\frac{G_m(x)}{g_{mQ}} \approx \frac{G_L}{g_{mQ}n}. \tag{6.4–6}$$

The second is that the total conductance G_T shunting the tuned circuit of Fig. 6.4–4 may be approximated by G_L. The total conductance G_T is given by

$$G_T = G_L + n^2 G_E + n^2 G_m(x)/\alpha\,. \tag{6.4–7}$$

With the substitution of $G_m(x)$ given by Eq. (6.4–4) into Eq. (6.4–7), G_T simplifies to

$$G_T = \frac{G_L + n^2 G_E}{1 - n/\alpha} \approx G_L, \tag{6.4–8}$$

which is the desired result. The implication here is that if $n \ll 1$ the total loss reflected through the transformer is negligible in comparison with G_L. Similarly,

$$Q_T = \frac{\omega_0 C}{G_T} \approx \frac{\omega_0 C}{G_L}. \tag{6.4–9}$$

In addition, for $n \ll 1$ it is readily demonstrated that

$$Q_E \approx Q_T \approx n Q_{T'}, \tag{6.4–10}$$

where

$$Q_E = \frac{\omega_0(C_1 + C_2)}{G_E + G_m(x)/\alpha} \quad \text{and} \quad nQ_{T'} = \frac{\omega_0 C}{n[G_E + G_m(x)/\alpha]};$$

hence $Q_T \gg 1$ and $n \ll 1$ ensures not only that $v_t(t)$ and $v_i(t)$ are sinusoidal but also that the capacitive transformer may be replaced by its ideal transformer model. [The reader should convince himself of the validity of Eq. (6.4–10).]

Example 6.4–1 For the Colpitts oscillator shown in Fig. 6.4–5, determine an expression for $v_o(t)$.

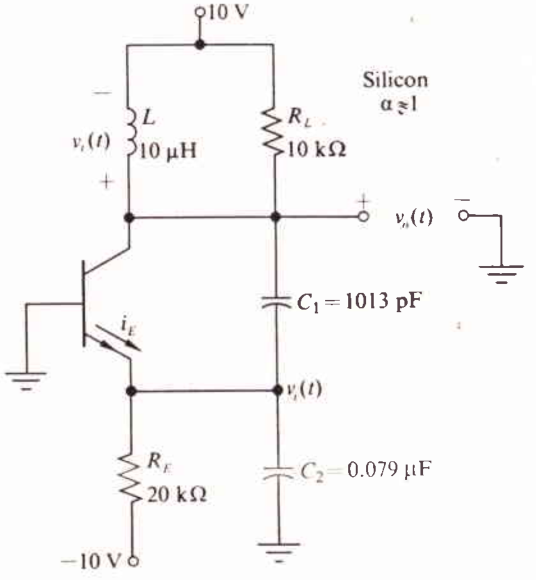


Figure 6.4–5

Solution. For the oscillator shown in Fig. 6.4–5 we note that

$$C = \frac{C_1 C_2}{C_1 + C_2} = 1000\text{ pF},$$

$$\omega_0 = \frac{1}{\sqrt{LC}} = 10^7\text{ rad/sec},$$

$$n = \frac{C_1}{C_1 + C_2} = \frac{1}{80},$$

$$V_\lambda = 9.3\text{ V},$$

$$I_{EQ} = \frac{9.3\text{ V}}{20\text{ k}\Omega} = 0.465\text{ mA},$$

$$g_{\text{in}\,Q} = \frac{qI_{EQ}}{kT} = \frac{1}{56\,\Omega} \approx g_{mQ},$$

and

$$Q_T \approx \omega_0 C R_L = 100.$$

Since $Q_T \gg 1$ and $n \ll 1$, $v_t(t)$ and $v_i(t)$ are sinusoidal and the model of Fig. 6.4–4 is valid; thus the oscillation grows at $\omega_0 = 10^7$ rad/sec and stabilizes at that value of x for which $G_m(x)/g_{mQ} = 0.448$ [cf. Eq. (6.4–4) or Eq. (6.4–6)]. Therefore, from Fig. 5.4–5 ($V_\lambda = 9.3\text{ V} \approx \infty$), we obtain $x \approx 3.5$ and in turn $V_1 = 99$ mV, $V_t = V_1/n = 7.9$ V, and finally

$$v_o = (10\text{ V}) + (7.9\text{ V})\cos 10^7 t,$$

which does not saturate the transistor.

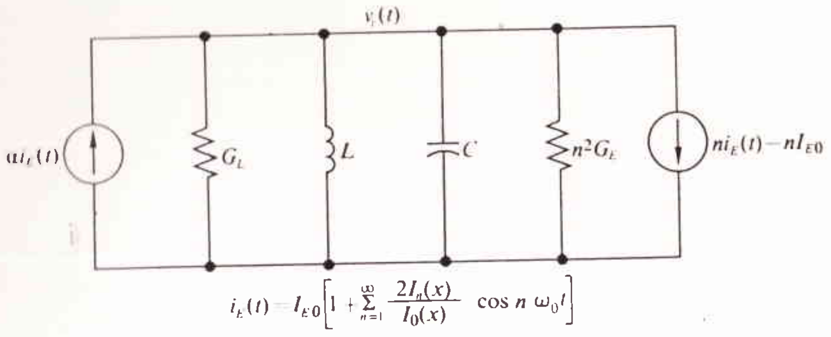
Total Harmonic Output Distortion

To obtain an expression for the total harmonic distortion (THD) of $v_t(t)$, we assume that the distortion of $v_t(t)$ and $v_i(t)$ is sufficiently small so that $i_E(t)$ in the circuit of Fig. 6.4–1 has the form

$$i_E(t) = I_{E0}\left[1 + \sum_{n=1}^{\infty} \frac{2I_n(x)}{I_0(x)} \cos \omega_0 t\right].$$

We then replace the transistor input by a current source $i_E(t)$ and reflect $i_E(t)$ through the capacitive transformer [by taking a Norton equivalent of the circuit to the right of G_L and assuming that G_E is negligible in comparison with $\omega(C_1 + C_2)$ over the frequency range, $\omega \geq \omega_0$, occupied by $i_E(t) - I_{E0}$] to obtain the circuit model shown in Fig. 6.4–6. For this circuit the amplitude of the fundamental component of $v_t(t)$ is given by

$$V_{t_1} = \frac{\alpha I_{E0}\left(1 - \dfrac{n}{\alpha}\right)\dfrac{2I_1(x)}{I_0(x)}}{G_L + n^2 G_E}, \tag{6.4–11}$$



$v_t(t)$
$\alpha i_e(t)$
G_L
L
C
n^2G_E
$ni_E(t) - nI_{E0}$
$i_E(t) = I_{E0}\left[1 + \sum_{n=1}^{\infty} \frac{2I_n(x)}{I_0(x)} \cos n\omega_0 t\right]$

Fig. 6.4–6 Circuit for obtaining THD of $v_t(t)$.

vhereas the amplitude of the kth harmonic component of $v_t(t)$ is given by

$$V_{tk} = \alpha I_{E0}\left(1 - \frac{n}{\alpha}\right)\frac{2I_k(x)}{I_0(x)}Z_{11}(jk\omega_0). \qquad (6.4\text{–}12)$$

f the Q_T of the circuit of Fig. 6.4–7 is high, then

$$Z_{11}(jk\omega_0) \approx \frac{k}{\omega_0 C(k^2 - 1)}$$

cf. Eq. 3.4–11). In addition, for $n \ll 1$,

$$G_L + n^2G_E \approx G_L;$$

herefore (cf. Section 3.5).

$$\text{THD} = \sqrt{\sum_{k=2}^{\infty}\left(\frac{V_{tk}}{V_{t1}}\right)^2} \approx \frac{G_L}{\omega_0 C}\sqrt{\sum_{k=2}^{\infty}\left(\frac{k}{k^2 - 1}\right)^2\left[\frac{I_k(x)}{I_1(x)}\right]^2} \approx \frac{D(x)}{Q_T}. \qquad (6.4\text{–}13)$$

vhere

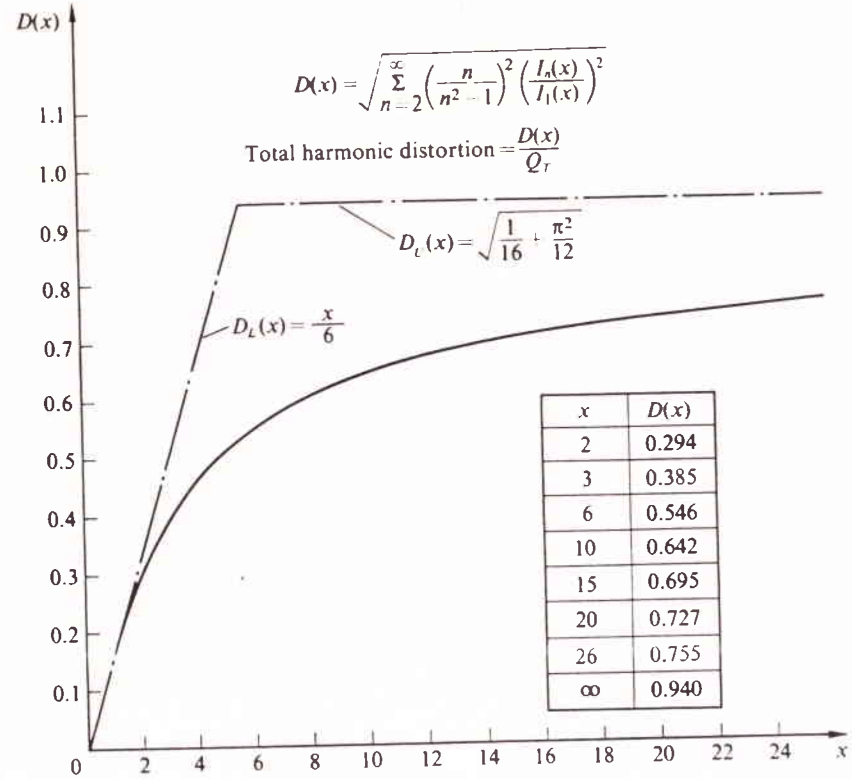
$$D(x) = \sqrt{\sum_{k=2}^{\infty}\left(\frac{k}{k^2 - 1}\right)^2\left[\frac{I_k(x)}{I_1(x)}\right]^2}.$$

A plot of $D(x)$ vs. x obtained by numerical evaluation of the significant terms of $D(x)$ is presented in Fig. 6.4–7. For large values of x,

$$\frac{I_k(x)}{I_1(x)} \to 1$$

and thus $D(x)$ approaches the upper asymptote

$$D_U(x) = \sqrt{\sum_{k=2}^{\infty}\left(\frac{k}{k^2 - 1}\right)^2} = \sqrt{\frac{1}{16} + \frac{\pi^2}{12}} = 0.94.$$



x	$D(x)$
2	0.294
3	0.385
6	0.546
10	0.642
15	0.695
20	0.727
26	0.755
∞	0.940

Fig. 6.4–7 Plot of $D(x)$ vs. x.

For small values of x,

$$\frac{I_2(x)}{I_1(x)} \approx \frac{x}{4} \gg \frac{I_k(x)}{I_1(x)} \qquad \text{for } k = 3, 4, \ldots$$

(cf. Appendix at back of book); therefore, $D(x)$ approaches the lower asymptote

$$D_L(x) = \frac{2}{3}\frac{I_2(x)}{I_1(x)} \approx \frac{x}{6}.$$

These asymptotes are also illustrated in Fig. 6.4–7.

From Fig. 6.4–7 it is apparent that, if the oscillator stabilizes with $x = 10$ ($V_1 = 260$ mV), then THD $= 0.642/Q_T$. Consequently Q_T must exceed 64.2 to keep the harmonic distortion below 1%. Clearly the smaller the value of x at which the oscillator stabilizes, the smaller the distortion. Unfortunately, a small steady-state value of x often results in poor amplitude stability, as we shall see in the following paragraphs.

Amplitude Stability

The amplitude stability of an oscillator is the sensitivity of the output oscillation amplitude V_t to variations in supply voltage, temperature, etc. In particular, the amplitude sensitivity factor is defined quantitatively as

$$S_\mu = \frac{\Delta V_t/V_t}{\Delta\mu/\mu}, \tag{6.4–14}$$

where μ is the parameter within the oscillator which is subject to variation, $\Delta\mu$ is the amount of the parameter variation, and ΔV_t is the corresponding variation in V_t. Small values of S_μ are desirable; the smaller the value of S_μ, the smaller the change in V_t with changes in μ.

With this definition we may evaluate the various amplitude stability factors for the oscillator of Fig. 6.4–1 by first writing the output voltage in the form

$$v_o(t) = -V_{CC} + V_t \cos\omega_0 t, \tag{6.4–15}$$

where $V_t = V_1/n = kTx/qn$ and x is the solution of the equation

$$\frac{G_m(x)}{g_{mQ}} = \frac{2I_1(x)}{xI_0(x)}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right] = \frac{G_L + n^2G_E}{n(1 - n/\alpha)g_{mQ}}.$$

We then vary the parameter of interest (μ) the desired amount ($\Delta\mu$) and determine the resultant value x from Eq. (6.4–5) or Fig. 5.4–5. This value of x immediately yields the new value of $V_t = kTx/qn$ from which ΔV_t may be computed and in turn S_μ evaluated.

Consider, for example, the evaluation of S_n for the oscillator of Example 6.4–1 as replacement of C_2 causes the transformer ratio n to increase from 1/80 to 3/160 (an increase of 50%—that is, $\Delta n/n = \frac{1}{2}$). With the new value of $n = 3/160$,

$$\frac{G_m(x)}{g_{mQ}} \approx \frac{G_L}{ng_{mQ}} = 0.298;$$

therefore, from Fig. 5.4–5 ($V_\lambda \approx \infty$) we obtain $x \approx 6.1$, from which it follows that $V_t = 8.45$ V, $\Delta V_t = 0.55$ V, and $S_n = 0.13$.

On the other hand, if G_L is increased from 100 μ℧ to 150 μ℧ in Example 6.4–1 (again a change of 50%), then $V_t = 4.7$ V, $\Delta V_t = 3.2$ V, and $S_{G_L} = -0.81$.

Several facts become apparent at this point. If we wish to minimize S_{G_L} or $S_{I_{EQ}} = S_{g_{mQ}}$, we must minimize the variations of x with G_L or I_{EQ}. We accomplish this by choosing the parameters of the oscillator so that the oscillation amplitude stabilizes in the vicinity of $x \approx 2$, at which value the slope of the $G_m(x)/g_{mQ}$ curve is steepest.

On the other hand, if we wish to minimize S_n for $n \ll 1$ and $V_\lambda \approx \infty$, we choose the parameters of the oscillator so that the oscillation amplitude stabilizes for the largest possible value of x such that $2I_1(x)/I_0(x) \approx 2$. For this condition,

$$\frac{G_m(x)}{g_{mQ}} \approx \frac{2}{x} \approx \frac{G_L}{ng_{mQ}}$$

and

$$V_t \approx \frac{2kTg_{mQ}}{qG_L} = \frac{2I_{EQ}}{G_L},$$

which is independent of n; thus $S_n \to 0$.

Similarly, if we assume that $g_{mQ} = qI_{EQ}/kT$ is the primary source of parameter variation with temperature, we obtain $S_T \to 0$ if the parameters of the oscillator are chosen so that the oscillation amplitude stabilizes for x as large as possible.

Let us now try to consolidate the above information in the form of an example dealing with the design of an oscillator.

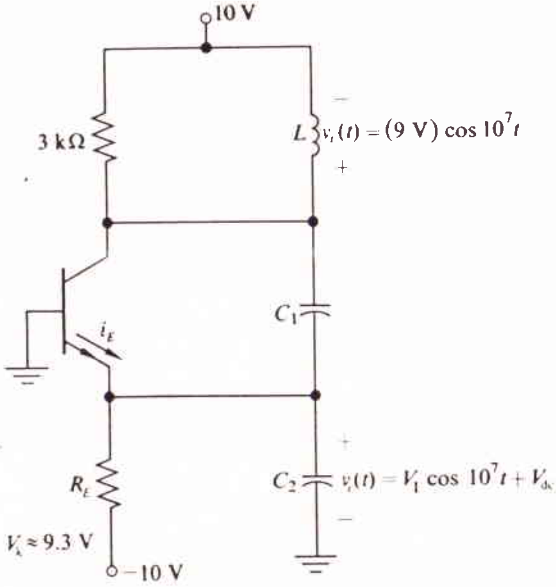

10 V
3 kΩ
L
$v_t(t) = (9\text{ V})\cos 10^7t$
i_E
C_1
R_E
C_2
$v(t) = V_1 \cos 10^7t + V_{dc}$
$V_\lambda \approx 9.3$ V
−10 V

Figure 6.4–8

Example 6.4–2 Design a Colpitts oscillator which provides an 18-V peak-to-peak, 10^7 rad/sec sine wave, having a total harmonic distortion of 1%, to a load resistor of 3 kΩ. The oscillator should make use of available ±10-V power supplies and provide reasonable first-order amplitude stability with temperature variations.

Solution. We first draw the schematic diagram of the Colpitts oscillator we intend to employ, as shown in Fig. 6.4–8. The design entails the choice of C_1, C_2, L, R_E, and the appropriate transistor to meet the required specifications.

To achieve good amplitude stability with temperature variations, we require the oscillator to stabilize with a large value of x or (V_1). On the other hand, a very large value of x requires a very large value of Q_T to achieve the desired THD; thus we compromise by designing the oscillator to stabilize for $x = 10$, which is achieved if (cf. Fig. 5.4–5 with $V_\lambda = 9.3$ V)

$$\frac{G_m(10)}{g_{mQ}} = \frac{G_L + n^2 G_E}{n g_{mQ}(1 - n/\alpha)} = 0.19. \tag{6.4–16}$$

With $x = 10$, from Fig. 6.4–8 we obtain $D(x) = 0.642$; thus to obtain THD = 0.01, we must have

$$Q_T \approx \omega_0 C R_L = 64.2,$$

or equivalently

$$C = \frac{C_1 C_2}{C_1 + C_2} = \frac{64.2}{(3\ \text{k}\Omega)(10^7\ \text{rad/sec})} = 2140\ \text{pF}.$$

Therefore, $L = 1/\omega_0^2 C = 4.67\ \mu$H. (Slug tuning of L is usually necessary to obtain the exact required value of ω_0.) In addition,

$$\frac{V_1}{V_t} = n = \frac{C_1}{C_1 + C_2} = \frac{(26\ \text{mV})(10)}{9\ \text{V}} = 0.029$$

(which ensures $V_t = 9$ V with $x = 10$). From the above values of C and n we obtain

$$C_2 = \frac{C}{n} = 0.074\ \mu\text{F} \qquad \text{and} \qquad C_1 = \frac{C}{1 - n} = 2200\ \text{pF}.$$

With a value for n, with G_L specified, and with the knowledge that $\alpha \approx 1$ and $n^2 G_E \ll G_L$, we finally observe from Eq. (6.4–16) that

$$g_{mQ} = \frac{qI_{EQ}}{kT} = \frac{q(V_\lambda/R_E)}{kT} = 0.0625\ ℧,$$

or equivalently,

$$R_E = 5.75\ \text{k}\Omega.$$

At first glance the choice of the dc bias resistor R_E might seem unimportant; however, this is not the case, since it directly determines I_{EQ}, which in turn sets the value of g_{mQ} and thus the steady-state operating point of the oscillator.

In choosing a transistor, two key factors must be considered: (1) the collector-to-base breakdown voltage and (2) the inherent frequency-limiting parameters of the transistor. In this example the maximum value of v_{CB} is 19 V; thus a transistor with a breakdown voltage in excess of 19 V must be chosen.

In general, if the transistor is chosen with an output capacitance small in comparison with C_1 and an input capacitance small in comparison with C_2, the transistor has little effect on the oscillator design. In general, low-power transistors can be found with sufficiently small values of input and output capacity so that designs similar to the one in this example are valid for oscillation frequencies in excess of 100 MHz.

Variable-Frequency Oscillator

If a single-transistor oscillator is to be used as a local oscillator or a local oscillator-mixer combination in a superheterodyne (AM, FM, or TV) receiver, the oscillator is usually constructed in the form shown in Fig. 6.4–9. The basic reason for this three-winding transformer configuration is to isolate the input and output capacity of the transistor from the tuning capacitor C by providing a step-down transformer turns ratio both to the collector and to the emitter from the L_2–C tuned circuit. Such an arrangement prevents changes in the transistor capacitance, which is quite sensitive to small variations in supply voltage, from affecting the frequency of oscillation, particularly at the high-frequency end of the tuning range where C is usually quite small.

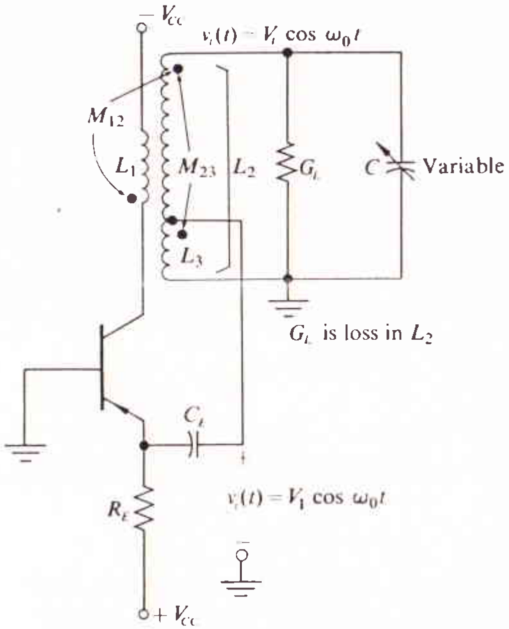


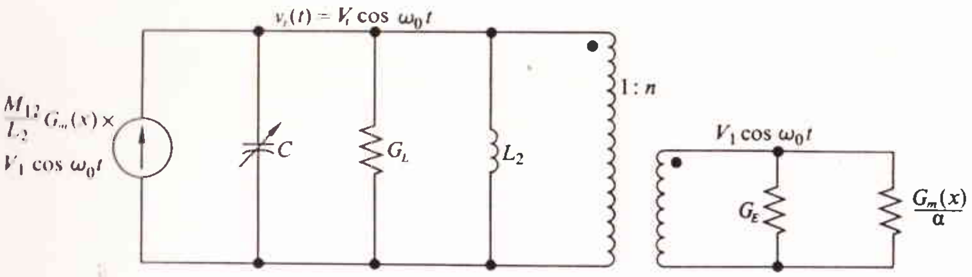
Fig. 6.4–9 Schematic diagram of local oscillator.

If the Q_T of the tuned circuit is assumed to be high, the oscillator of Fig. 6.4–9 has the large-signal model for determining $v_t(t)$ as shown in Fig. 6.4–10; therefore, the oscillation has a radian frequency of $\omega_0 = 1/\sqrt{L_2 C}$ and stabilizes at that value of $x = qV_1/kT$ for which

$$\frac{(M_{12}/L_2)nG_m(x)}{G_L + n^2[G_m(x)/\alpha + G_E]} = 1, \tag{6.4–17}$$

or equivalently,

$$\frac{G_m(x)}{g_{mQ}} = \frac{G_L + n^2 G_E}{n g_{mQ}(M_{12}/L_2 - n/\alpha)}. \tag{6.4–18}$$



$$\omega_0 = 1/\sqrt{L_2C}, \quad n = M_{23}/L_2$$

Fig. 6.4–10 Large-signal local oscillator model.

Again x is determined from Fig. 5.4–5, and in turn V_t is obtained as $V_t = kTx/qn$. Although $v_t(t)$ and $v_i(t)$ are reasonably sinusoidal, we should recall that the collector voltage is not sinusoidal because of the presence of the series inductor $(1 - k_{12}^2)L_1$ (cf. Section 2.5).

Collector Saturation

If the value calculated for the amplitude of the fundamental component of collector current I_{C1} on the assumption of no collector saturation exceeds $V_{tS}/R_T = (V_{CC} + 0.7)/R_T$ (in Fig. 6.4–1), then collector saturation does occur and the tuned-circuit voltage is limited to $V_{CC} + 0.7$. In this case the emitter-base driving voltage is n times the known tuned-circuit voltage and the amplitude of the fundamental collector current is

$$I_{C1} = G_m(x)V_t n = \alpha I_{E0} \frac{2I_1(x)}{I_0(x)},$$

where $x = qn(V_{CC} + 0.7)/kT$ in the saturation case. The difference between this current and the allowable net fundamental current into the tuned circuit is the component removed by the collector-base saturation pulse train.

It is apparent that, if the loaded Q_T of the tuned circuit is high, the effect of collector saturation may be represented as an equivalent linear conductance shunting the tuned circuit to yield a total shunt conductance G_{TS} given by $G_{TS} = V_{tS}/I_{C1}$, which for the circuit of Fig. 6.4–1 reduces to

$$G_{TS} = \frac{\alpha I_{E0}}{V_{CC} + 0.7} \frac{2I_1(x)}{I_0(x)}. \tag{6.4–19}$$

Note that this equation is valid when $G_{TS} > G_T$, that is, only so long as saturation does occur on the peak of each cycle.

Thus, if collector saturation occurs, the output amplitude is directly proportional to V_{CC}, while the effective loaded Q of the circuit depends on V_{CC}, I_{EQ}, and n. Specifically, $Q_T = \omega_0 C/G_{TS}$.

Example 6.4–3 For the Colpitts oscillator shown in Fig. 6.4–11, determine an expression for $v_o(t)$.

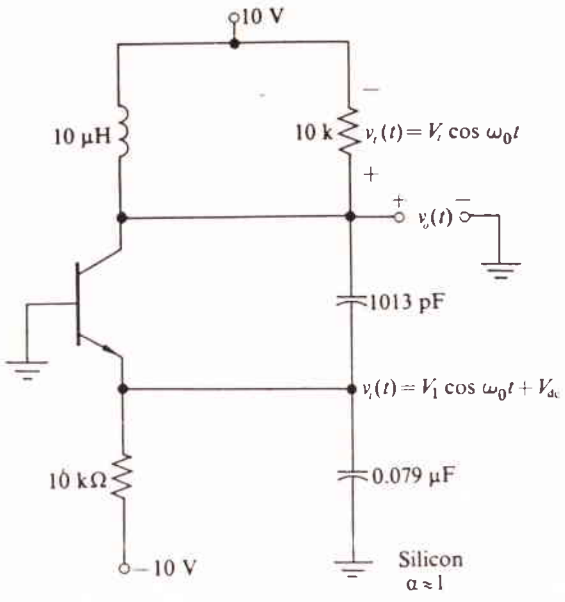


Figure 6.4–11

Solution. For the oscillator shown in Fig. 6.4–11, we note that

$$C = \frac{C_1C_2}{C_1 + C_2} = 1000 \text{ pF},$$

$$\omega_0 = \frac{1}{\sqrt{LC}} = 10^7 \text{ rad/sec},$$

$$n = \frac{C_1}{C_1 + C_2} = \frac{1}{80},$$

$$V_\lambda = 9.3 \text{ V},$$

$$I_{EQ} = 0.93 \text{ mA},$$

and

$$g_{mQ} = 1/28\ \Omega.$$

identical to the large-signal model of Example 6.4–3 (except for two compensating phase inversions), we conclude that oscillations occur at a radian frequency of $\omega_0 = 1/\sqrt{L_1 C}$ and stabilize at that value of x for which[†]

$$\frac{G_m(x)}{g_m} = \frac{4a_1(x)}{x} = \frac{G_L}{g_m n(1 - n/\beta)} \approx \frac{G_L}{n g_m}. \tag{6.5–1}$$

In this case, x is found by plotting $G_L/g_m n(1 - n/\beta)$ on the curve of Fig. 4.6–5. With the resultant value of x, we obtain $V_1 = kTx/q$ and $V_t = V_1/n$.

For the self-limiting differential-pair oscillator, the total harmonic distortion (THD) is readily found, by methods similar to those used for the single-transistor oscillator, to be of the form

$$\text{THD} = \frac{D(x)}{Q_{TU}} \approx \frac{D(x)}{Q_T}, \tag{6.5–2}$$

where $Q_{TU} = \omega_0 C/G_L$ and $D(x)$ is given by

$$D(x) = \sqrt{\sum_{n=2}^{\infty} \left[\frac{2n - 1}{(2n - 1)^2 - 1}\right]^2 \left[\frac{a_{2n-1}(x)}{a_1(x)}\right]^2}. \tag{6.5–3}$$

The Fourier coefficients $a_n(x)$ are defined by Eq. (4.6–10). A plot of $D(x)$ vs. x obtained by numerical evaluation of the significant terms of $D(x)$ is presented in Fig. 6.5–3. Note that, for a given value of x, $D(x)$ for the differential-pair oscillator is lower, by more than a factor of 7, than the corresponding $D(x)$ for the single-transistor oscillator. This, of course, is an anticipated result.

The upper asymptote $D_U(x)$ shown in Fig. 6.5–3 is obtained by noting that as $x \to \infty$, i_{E1} and i_{E2} approach square waves and thus

$$\left[\frac{a_{2n-1}(x)}{a_1(x)}\right]^2 \to \frac{1}{(2n - 1)^2}.$$

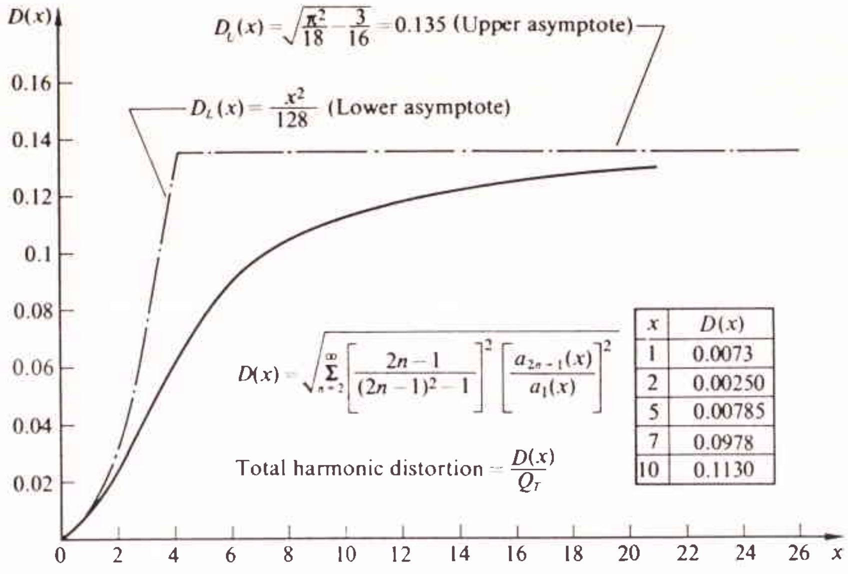
Therefore,

$$D_U(x) = \sqrt{\sum_{n=2}^{\infty} \left[\frac{1}{(2n - 1)^2 - 1}\right]^2} = \sqrt{\frac{\pi^2}{18} - \frac{3}{16}} = 0.135. \tag{6.5–4}$$

For small values of x,

$$\left.\frac{a_3(x)}{a_1(x)}\right| \approx \frac{x^2}{48} \gg \frac{a_{2n-1}(x)}{a_1(x)} \qquad \text{for } n = 3, 4, 5, \ldots$$

[†] The approximation of $G_L/g_m n(1 - n/\beta)$ by $G_L/g_m n$ is valid even if n is not much less than 1, provided that β is large. This is a direct result of the fact that the base, rather than the emitter, is being driven and thus the reflected loading across G_L is reduced by a factor of β from that obtained with a corresponding emitter drive. Therefore, in the differential pair, $Q_T \approx \omega_0 C/G_L \gg 1$ and $\beta \gg 1$ replace the conditions of $Q_T \gg 1$ and $n \ll 1$ required of a single transistor as sufficient conditions to ensure the validity of the ideal transformer model, to ensure the validity of assuming $v_i(t)$ to be sinusoidal, and to ensure $G_T \approx G_L$.



x	$D(x)$
1	0.0073
2	0.00250
5	0.00785
7	0.0978
10	0.1130

Fig. 6.5–3 Plot of $D(x)$ vs. x for the differential pair.

(cf. Section 4.6); therefore, $D(x)$ approaches the lower asymptote

$$D_L(x) = \frac{3}{8}\frac{x^2}{48} = \frac{x^2}{128}, \tag{6.5–5}$$

which is also plotted in Fig. 6.5–3.

Example 6.5–1 For the oscillator shown in Fig. 6.5–4, determine an expression for $v_o(t)$. Also determine the THD of $v_o(t)$.

Solution. For the circuit of Fig. 6.5–4,

$$n = \frac{M}{L_1} = 0.02 \ll 1, \qquad \omega_0 = 10^7 \text{ rad/sec},$$

and

$$Q_T \approx Q_{TU} = \omega_0 C R_L = 50;$$

therefore, if we assume that neither transistor saturates, then the large-signal model of Fig. 6.5–2 may be employed. With the use of the model we conclude that the oscillation frequency is 10^7 rad/sec and that the oscillation amplitude stabilizes for that value of $x = qV_1/kT$ for which

$$\frac{G_m(x)}{g_m} = \frac{4a_1(x)}{x} = \frac{G_L}{n(1 - n/\beta)g_m} \approx \frac{G_L}{n g_m}.$$

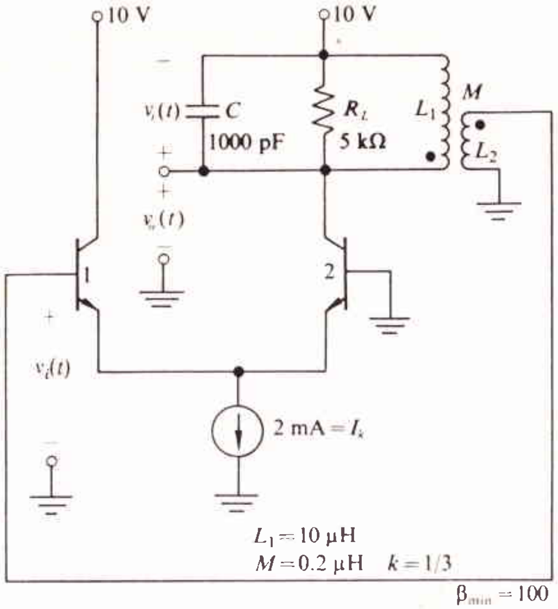


Figure 6.5–4

For the circuit of Fig. 6.5–4,

$$g_m = \alpha \frac{g_{\text{in}}}{2},$$

where

$$g_{\text{in}} = \frac{qI_{EQ}}{kT} = \frac{1}{2}\frac{qI_k}{kT} = \frac{1}{26\ \Omega};$$

hence $G_m(x)/g_m = 0.52$, which, with the aid of Fig. 4.6–5, yields $x = 4.2$ and in turn

$$V_t = \frac{(26\text{ mV})(4.2)}{0.02} = 5.45\text{ V}.$$

Consequently,

$$v_o(t) = (10\text{ V}) + (5.45\text{ V})\cos 10^7 t;$$

this value of v_o does not saturate the collector of transistor 2.

Since $x = 4.2$, from Fig. 6.5–3 we obtain $D(x) = 0.065$ and in turn

$$\text{THD} = \frac{0.065}{50} = 0.0013 \quad (0.13\%);$$

this value is quite small, as expected.

6.6 SELF-LIMITING JUNCTION FIELD EFFECT TRANSISTOR OSCILLATORS

Like the differential pair, the junction field effect transistor may be used as the active element in any of the self-limiting oscillators discussed in Section 6.4. In addition, the FET may be employed as the active element of the tuned-gate oscillator shown in Fig. 6.6–1.

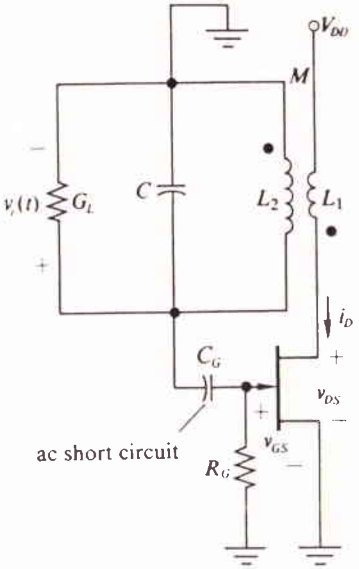


Fig. 6.6–1 Tuned-gate junction FET oscillator.

The tuned-gate oscillator is probably the most logical configuration for a self-limiting FET oscillator, since in this configuration the tuned circuit is virtually not loaded at all by the FET and therefore it is possible to maintain a high value of Q_T with its accompanying good frequency stability. Obviously R_G, which is usually chosen anywhere between 1 MΩ and 10 MΩ, does not appreciably load the tuned circuit. In addition, with $M/L_2 \ll 1$, only an exceptionally small amount of loading due to the output impedance (both capacitive and resistive) of the FET is reflected through the transformer. The ac component of v_{DS} is also kept small with $M/L_2 \ll 1$, the result being a negligibly small "Miller effect"; therefore, the capacitive loading of the FET input directly across the tuned circuit is minimized. The Colpitts, Hartley, and tuned-drain configurations possess none of these advantages and therefore do not find as wide application as does the oscillator shown in Fig. 6.6–1.

Note that the FET oscillator of Fig. 6.6–1 is clamp biased; i.e., the combination of the capacitor C_G (which is an ac short circuit at the oscillation frequency) and the gate-to-source junction diode clamps the positive peak of the gate-to-source voltage essentially to zero.† Thus if $v_i = V_1 \cos \omega_0 t$, then $v_{GS} = V_1(\cos \omega_0 t - 1)$. This form f bias, in contrast with that obtained by inserting a negative voltage source in series vith R_G, thus applying a fixed gate-to-source bias, has the advantage that the oscilla-ion amplitude may be stabilized with the FET operating completely within its quare-law region. This property is quite desirable when the circuit of Fig. 6.6–1 is mployed as an "oscillating mixer" (Chapter 7), in which case square-law operation ninimizes undesired intermodulation products. With a clamped bias, the FET ransconductance decreases with increasing sinusoidal input voltage amplitude vithin the square-law region (a necessary condition for amplitude stabilization—cf. ig. 4.9–4); with a fixed bias, on the other hand, the FET transconductance remains ndependent of input voltage amplitude for operation completely within the square-aw region (cf. Eq. 4.4–6).

---

† Actually v_{GS} is clamped to V_0, the turn-on bias of the diode. Since V_0 is a function of the average diode current (cf. Fig. 5.1–6) and since an R_G of several megohms requires the average diode current to be quite small ($\frac{1}{10}$ μA or less), V_0 is usually required to be less than $\frac{1}{2}$ V, even for silicon; thus as a good first approximation we may usually assume $V_0 \approx 0$.

Note also that the transformer in Fig. 6.6–1 with $M/L_2 \ll 1$ keeps the reflected mpedance in the drain cricuit quite small; hence the output resistance r_o of the FET nay be neglected in obtaining the FET model, or equivalently the drain current i_D nay be expressed in the form

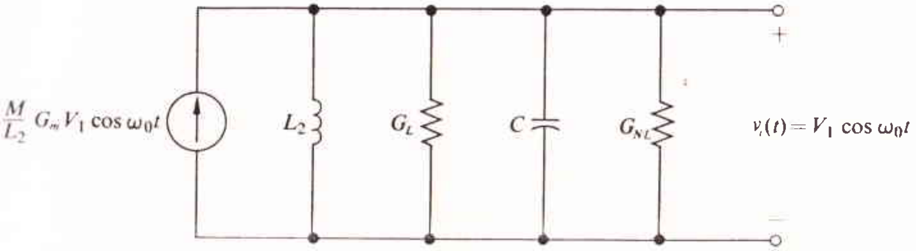
$$i_D = I_{DSS}\left(1 - \frac{v_{GS}}{V_p}\right)^2, \tag{6.6–1}$$

here I_{DSS} is the value of i_D when $v_{GS} = 0$ and $v_{DS} = V_{DD}$, and V_p is the pinch-off oltage of the FET. In addition, $M/L_2 \ll 1$ ensures that the ac drain voltage is quite mall, causing, in general, operation completely within the saturation region of the ET.

From the above considerations and the assumption that Q_T of the tuned circuit sufficiently high to keep $v_i(t)$ sinusoidal (that is, $v_i(t) = V_1 \cos \omega_0 t$), we may obtain he large-signal model shown in Fig. 6.6–2 for the oscillator of Fig. 6.6–1. In this nodel, because of the assumed high Q_T, only the fundamental component of drain urrent is reflected through the transformer, in the fashion of Fig. 2.5–7, to obtain the riving current source

$$\frac{M}{L_2} G_m V_1 \cos \omega_0 t,$$

here G_m is the large-signal fundamental transconductance of the clamp-biased ET. (A plot of G_m/g_{m0} vs. $-V_1/V_p$ is shown in Fig. 4.9–4.) In addition, because of


$\frac{M}{L_2} G_m V_1 \cos \omega_0 t$
L_2
G_L
C
G_{NL}
+
$v_i(t) = V_1 \cos \omega_0 t$
−

$$\omega_0 = \frac{1}{\sqrt{L_2 C}}, \quad G_{NL} = \frac{3}{R_G}, \quad Q_T = \frac{\omega_0 C}{G_{NL} + G_L}$$

Fig. 6.6–2 Large-signal model for the tuned-gate FET oscillator.

the assumption of high Q_T, the clamp-biasing circuit is replaced by its equivalent conductance $G_{NL} = 3/R_G$, which is developed in Example 5.5–5.

With the model of Fig. 6.6–2, we observe that

$$\mathscr{I}m\, A_L(j\omega_0) = 0 \qquad \text{for} \qquad \omega_0 = \frac{1}{\sqrt{L_2 C}}$$

and that

$$\mathscr{R}e\, A_L(j\omega_0) = \frac{M}{L_2} \frac{G_m}{G_L + 3G_G} = 1 \tag{6.6–2}$$

for that value of V_1 for which

$$\frac{G_m}{g_{m0}} = \frac{L_2}{M} \frac{G_L + 3G_G}{g_{m0}}, \tag{6.6–3}$$

where $g_{m0} = 2I_{DSS}/-V_p$ is the small-signal transconductance of the FET with $v_{GS} = 0$. Clearly, once V_p, I_{DSS}, G_L, and G_G are specified, $-V_1/V_p$, and in turn V_1, may be determined from Fig. 4.9–4.

As an example, consider

$$L_2 = 10\,\mu\text{H}, \qquad M = 1\,\mu\text{H}, \qquad C = 10\,\text{pF} = C_G,$$
$$R_G = 3\,\text{M}\Omega, \qquad R_L = 50\,\text{k}\Omega, \qquad V_p = -4\,\text{V},$$

and

$$I_{DSS} = 4\,\text{mA}.$$

Therefore,

$$\omega_0 = 10^8\,\text{rad/sec}, \qquad Q_T = 47.5 > 10,$$
$$g_{m0} = 2000\,\mu\text{mho}, \qquad \text{and} \qquad \frac{G_m}{g_{m0}} = 0.105.$$

From Fig. 4.9–4 we obtain $-V_1/V_p = 1.67$; hence $V_1 = 6.67$ V and

$$v_i(t) = (6.67\,\text{V}) \cos 10^8 t.$$

6.7 CRYSTAL OSCILLATORS

Since 1921 there have been hundreds of articles, tens of chapters in textbooks, and several books about crystal oscillators and piezoelectric resonators. Gerber and Sykes† have provided a recent summary of the state of the art with respect to crystals, while Hafner‡ has written extensively on crystal unit models and the measurement nodel parameters. Sections 6.12 and 6.13 and Chapter 9 of Edson† and Sections 8.2, and 8.3 of Groszkowski‡ provide textbook summaries of many quartz crystal perties and also discuss vacuum-tube crystal oscillator circuits in detail.

---

† E. A. Gerber and R. A. Sykes, "State of the Art—Quartz Crystal Units and Oscillators," *Proc. IEEE*, pp. 103–116 (Feb. 1966). Summary of data and characteristics; 63-item bibliography.

‡ E. Hafner, "The Piezoelectric Crystal Unit—Definitions and Methods of Measurement," *Proc. IEEE*, pp. 179–201 (Feb. 1969). Includes a 46-item bibliography.

It is not our purpose here either to provide design information about piezotric resonators or to catalog all possible crystal oscillator circuits. We wish ply to point out the basic properties of these resonators and to explain how they y be exploited to produce satisfactory circuits.

The reason why one employs a piezoelectric resonator in place of a conventional C combination is that the available Q in these mechanical vibration devices may up to 1000 times greater than that available with conventional elements. As we e seen, indirect frequency stability is directly proportional to Q; hence crystal illators are used as fixed-frequency, highly stable frequency or timing sources.

The principal material used as a mechanical resonator for oscillators is crystalline rtz. The properties of this material differ in different directions through the stal; in addition, each different "cut" may be mounted and vibrated in several erent ways. These different "cuts" have different temperature properties, different ges of possible operation, and different relations between their dimensions and ir mechanical and electrical properties.

Table 6.5 in Edson summarizes the properties of twelve of the more useful "cuts." me "cuts," for example the GT cut, have frequency shifts of considerably less n one part per million per degree Celsius over a 100°C range. Other "cuts" ve less desirable temperature properties but may have higher Q's, lower series istances, or more desirable properties at harmonics. The basic electrical model a properly mounted quartz crystal and its holder is shown in Fig. 6.7–1.

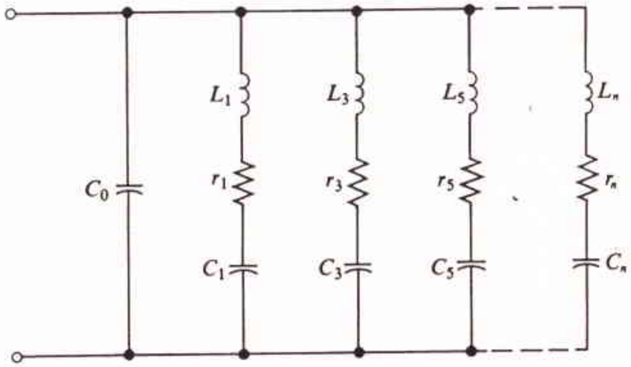
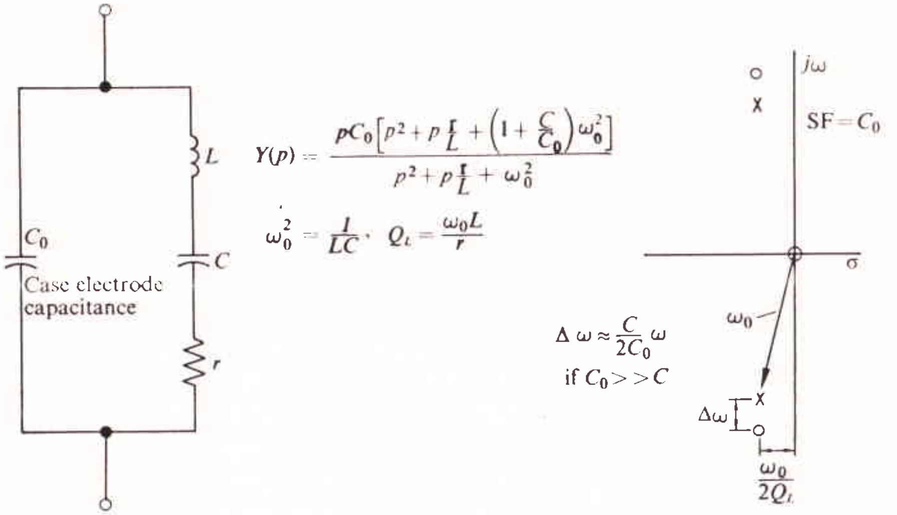


Fig. 6.7–1 Electrical model for a quartz crystal resonator.

As we shall see, the crystal is normally operated within 1% of the series resonant quency of one of the shunt branches; hence the circuit normally is reduced to C_0 shunt with a single series resonant circuit. The multiple branches result from mechanical vibrations at approximately the odd harmonics (usually called overtones) of the fundamental frequency. Since the mechanical frequency of the fundamental vibration is proportional to the crystal dimensions, practical considerations usually limit fundamental operation to the neighborhood of 20 MHz or less. By operating on an "overtone" frequency, crystal oscillators up to the neighborhood of 200 MHz are possible. For higher frequencies we must employ harmonic multiplication from a lower-frequency circuit.

If operation at a particular overtone is desired, then this should be specified when ordering the crystal. If a "cut" and mounting are used that minimize overtone effects, then sinusoidal operation may be possible in circuits with no filtering elements except the crystal itself.



L
C_0
Case electrode capacitance
C
r
$Y(p) = \frac{pC_0\left[p^2 + p\frac{r}{L} + \left(1 + \frac{C}{C_0}\right)\omega_0^2\right]}{p^2 + p\frac{r}{L} + \omega_0^2}$
$\omega_0^2 = \frac{1}{LC}, \quad Q_L = \frac{\omega_0 L}{r}$
$\Delta\omega \approx \frac{C}{2C_0}\omega$
if $C_0 >> C$
$j\omega$
$\text{SF} = C_0$
σ
ω_0
$\Delta\omega$
$\frac{\omega_0}{2Q_L}$

Fig. 6.7–2 Model for a crystal, its admittance equation, and its pole-zero diagram.

Figure 6.7–2 repeats the simplified version of the model together with its admittance relationship and pole-zero pattern.

In order to understand the particular properties of crystals it is useful to consider typical numerical values for the crystal parameters, for example,

$$\omega_0 = 10^7 \text{ rad/sec},$$

$$C/C_0 = 1/100,$$

$$C_0 = 4 \text{ pF},$$

and

$$Q_L = 20{,}000,$$

which lead to $C = 0.04$ pF, $L = 250$ mH, and $r = 125\ \Omega$. For this typical case, the vertical spacing between the pole and the zero in the vicinity of ω_0 is approximately

---

/. A. Edson, *Vacuum Tube Oscillators.* John Wiley, New York (1953).

Groszkowski, *Frequency of Self Oscillations.* Macmillan, New York (1964).

$_0/200$, while the horizontal distance from the j-axis to a complex pole or to a omplex zero is $\omega_0/40{,}000$, or 1/200 times the vertical spacing between the complex ole and zero locations. Thus near a complex pole one has in effect an isolated very gh-Q pole, while near a complex zero one has in effect an isolated very high-Q zero.

Two choices remain before we can select a circuit configuration. We must ecide whether we want to use the crystal as a low impedance [near a zero in $Z(p)$ a pole in $Y(p)$] or as a high impedance [near a pole in $Z(p)$ or a zero in $Y(p)$], nd we have to choose a method of amplitude limiting for the circuit.

Series resonant modes of operation usually place the crystal directly in the feed-ack path or use it to shunt a bias resistor. In both cases, the object is to keep the oop gain below the minimum required for oscillation except at the series resonant equency of the crystal. *Since the pole of $Y(p)$ [zero of $Z(p)$] is independent of C_0 ich circuits should not be affected by circuit-caused variations in the crystal's shunt apacitance.*

A capacitance C_x in series with the crystal modifies the scale factor of $Y(p)$ from $_0$ to

$$\frac{C_0 C_x}{C_0 + C_x}$$

nd shifts the poles of $Y(p)$ from ω_0 to

$$\sqrt{\frac{C + C_0 + C_x}{C_0 + C_x}}\,\omega_0.$$

hen $C_0 \gg C$, this capacitance changes the approximate vertical pole-zero spacing om $C\omega_0/2C_0$ to

$$\frac{C\omega_0}{2C_0}\frac{C_x}{C_0 + C_x}.$$

hus the possibility of a slight adjustment of the series resonant frequency is purchased the expense of a narrowing of the pole-zero spacing.

If $C_0 = 100C$ and C_x varies from C_0 to $C_0/5$, it is possible to have a frequency ning of 0.166% of ω_0 while reducing the minimum pole-zero spacing by only $\frac{1}{2}$.

A great many vacuum-tube oscillators employ the crystal operating in a high-npedance mode. Field effect transistor oscillators tend to follow this same pattern. all these circuits the operating frequency tends to be below the zero of $Y(p)$ [pole $Z(p)$] at a distance ω_0/Q_L or less from it. At these frequencies the crystal looks ke a parallel inductance-resistance combination and is used to replace an inductance a Colpitts-like or Hartley-like configuration. Such arrangements can be "tuned" varying the shunt capacitance, C_0.

For operation at higher frequencies and higher overtones, the pole-zero spacing ay be increased by adding a shunt inductance L_S that resonates with C_0 at the perating frequency. With such an arrangement a new pair of complex poles in (p) is formed; these poles lie between the complex zeros of $Z(p)$ and the origin. For

---

$C_0 L_S = LC$ and $C_0 \gg C$, the approximate vertical spacings of the two complex poles about the unchanged zero position is

$$\left(\pm\frac{1}{2}\sqrt{\frac{C}{C_0}} + \frac{C}{4C_0}\right)\omega_0.$$

which for $C_0 = 100C$ leads to $+(21/400)\omega_0$ and $-(19/400)\omega_0$. We should note that these spacings are ten times the spacing for the unshunted case. Since C/C_0 tends to decrease as the square of the overtone multiple, the shunt inductance allows us to keep a nearly constant percentage pole-zero spacing at the higher frequencies.

Varying C_0 (or the shunt inductance) provides a means of varying both of the complex poles. Operation of the crystal as an "inductance" is possible just below either pole, but is not in general desirable, since the pole positions depend strongly on the external shunt inductance.

A well-designed crystal oscillator circuit should stop oscillating if the crystal is removed and should minimize the crystal driving voltage and current. The last requirement is desirable not only to avoid physical damage through overdriving, but also to minimize temperature shifts due to self-heating. For example, if the thermal resistance of the mounted unit is 500°C/W, then an internal dissipation of 2 mW leads to a 1°C temperature rise; this temperature rise may lead, in some crystal "cuts" at some ambient temperatures, to frequency shifts of several parts per million.

Most very high-quality crystal oscillator circuits separate the frequency generation and the amplitude limiting functions by employing a separate control mechanism. Section 6.9 discusses several such circuits. In this section we concentrate on simple self-limiting circuits, beginning with the use of a single bipolar transistor as both the amplifying and the limiting device.

As an example of a single-transistor series mode circuit we consider adding a crystal between $v_i(t)$ and the emitter in the circuit of Example 6.4–1, as shown in Fig. 6.7–3. Here the tuned circuit is tuned to the pole of the crystal admittance $Y(p)$ to ensure operation at the correct overtone. Clearly the loop gain will be insufficient for operation at all other overtones because of the low tuned-circuit impedance.

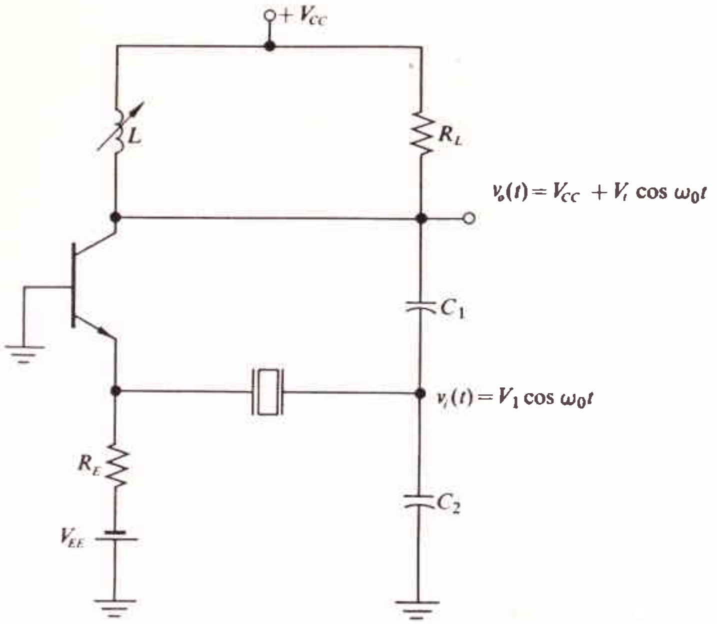
Intuitively we see that, if the crystal is series resonant and acts like a resistor, then this resistor and the transistor input act like an attenuator and reduce the loop gain. Hence it is likely that, if the circuit is to operate at all, we will have to restore the loop gain either by increasing n or by increasing the load resistor.

If the crystal acts merely as a resistor, then we could calculate the new operating amplitude using Eq. (6.4–3) and Fig. 4.8–6. However, since the crystal is acting like a high-Q series resonant circuit, it restricts the current through itself to a fundamental component.† If we assume a current source drive of I_{EQ} for the dc transistor bias, then this ac crystal current is equal to the ac emitter current of the transistor; that is,

$$I_{E1} = \frac{V_1 - V_{EB1}}{r}, \tag{6.7–1}$$

---

† This is the dual of a high-Q parallel resonant circuit where only the fundamental voltage exists.


+ V_{CC}
L
R_L
$v_o(t) = V_{CC} + V_i \cos \omega_0 t$
C_1
$v_i(t) = V_1 \cos \omega_0 t$
R_E
V_{EE}
C_2

r = series-resonant resistance of crystal,

$\omega_0 = \dfrac{1}{\sqrt{LC}}$ = series resonance of crystal,

$n = C_1/(C_1 + C_2)$, $C = C_1 C_2/(C_1 + C_2)$

Fig. 6.7–3 Series mode crystal oscillator.

where V_{EB1} is the amplitude of the fundamental emitter-base voltage. Now V_{EB1} is a function of I_{E1} and I_{EQ} which can be found by assuming that an ac current source and a dc current source drive the emitter-base junction and calculating the fundamental component of the emitter voltage. [Note that the emitter-base voltage is no longer sinusoidal, since the crystal separates the junction from the assumed sinusoidal voltage $v_i(t)$.]

If

$$i_E = I_{ES} e^{v_{EB} q/kT} = I_{EQ} + I_{E1} \cos \omega t, \tag{6.7–2}$$

then

$$v_{EB} = \frac{kT}{q} \ln \frac{I_{EQ}}{I_{ES}} + \frac{kT}{q} \ln \left(1 + \frac{I_{E1}}{I_{EQ}} \cos \omega t\right). \tag{6.7–3}$$

Hence, if the last term in Eq. (6.7–3) is expanded in a Fourier series, we may determine V_{EB1} to be of the form

$$V_{EB1} = \frac{kT}{q}\left[\frac{2}{y}\left(1 - \sqrt{1 - y^2}\right)\right], \tag{6.7–4}$$

where $y = I_{E1}/I_{EQ}$; in addition,

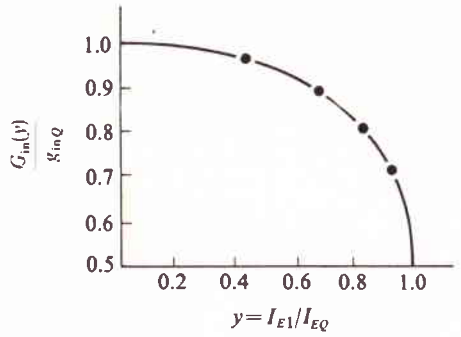
$$V_{EB2} = \frac{kT}{q}\left[\frac{1}{y^2}\left(y^2 - 2 + 2\sqrt{1 - y^2}\right)\right]. \tag{6.7–5}$$

For this case we define a large-signal average fundamental transistor input conductance as

$$G_{\text{in}}(y) = \frac{I_{E1}}{V_{EB1}} = \frac{I_{E1}}{I_{EQ}} \cdot \frac{I_{EQ}}{V_{EB1}} = \frac{qI_{EQ}}{kT} \frac{y^2}{2\left(1 - \sqrt{1 - y^2}\right)}$$

$$= g_{\text{in}\,Q} \frac{y^2}{2\left(1 - \sqrt{1 - y^2}\right)} \approx g_{\text{in}\,Q}\left(1 - \frac{y^2}{4}\right)$$

for $y \leq 0.6$. A plot of $G_{\text{in}}(y)/g_{\text{in}\,Q}$ is shown in Fig. 6.7–4. If we now assume that the reflected loading of the crystal resistance r in series with the transistor input resistance



y	$G_{\text{in}}/g_{\text{in}\,Q}$
0.0	1.000
0.2	0.990
0.3	0.977
0.4	0.961
0.5	0.937
0.6	0.900
0.8	0.800
0.9	0.725
1.0	0.500

Fig. 6.7–4 Normalized input conductance for a current-driven junction.

$1/G_{\text{in}}(y)$ across R_L is negligible, then the loop gain for the circuit of Fig. 6.4–3 may be written as

$$A_L(j\omega_0) = \frac{\alpha n R_L}{r + R_{\text{in}}(y)}; \tag{6.7–6}$$

hence the oscillation stabilizes for that value of $y = I_{E1}/I_{EQ}$ for which $A_L(j\omega_0) = 1$, or equivalently,

$$\frac{G_{\text{in}}(y)}{g_{\text{in}\,Q}} = \frac{1}{(\alpha n R_L - r) g_{\text{in}\,Q}}, \tag{6.7–7}$$

where r is the series resistance of the crystal, R_L is the load resistance, and $n = C_1/(C_1 + C_2)$.

It is apparent from Fig. 6.4–4 that for reasonable amplitude stability I_{E1} should between 0.4 and 1 times I_{EQ} or that

$$0.50 \leq \frac{1}{g_{\text{in}Q}(\alpha n R_L - r)} \leq 0.95. \tag{6.7–8}$$

In order to calculate the operating amplitude, one calculates $G_{\text{in}}(y)/g_{\text{in}Q}$ from q. (6.7–7) and uses Fig. 6.7–4 to find y. Then, since $V_t = \alpha I_{E1} R_L$, one finds that

$$V_t = \alpha y I_{EQ} R_L.$$

As an example, we might modify the circuit of Fig. 6.4–5 to include the crystal, in Fig. 6.7–3. If we increase n from 1/80 to 1/50 while maintaining all the other antities at their previous values, then for $\alpha = 0.98$ and $r = 125\ \Omega$,

$$\frac{G_{\text{in}}(y)}{g_{\text{in}Q}} = \frac{56}{71} = 0.79;$$

o, from Fig. 6.7–4, $y = 0.82$, so that $I_{E1} = 0.82 I_{EQ}$ and $V_t = 3.7$ V or

$$v_o(t) = (10\ \text{V}) + 3.7\ \text{V} \cos \omega t.$$

e effective loaded Q of the crystal is $\omega_0 L$ divided by $r + R_{\text{in}} = 125 + 71 \approx 196\ \Omega$; nce the crystal Q has been reduced from 20,000 to 12,800.

The crystal dissipation for this case is $I_{E1}^2 r/2 = 8.5\ \mu W$; this amount of dissipation ll certainly not cause excessive frequency shifts.

In making the previous calculations we assumed that the tuned circuit consisting L, R_L, and $C_1C_2/(C_1 + C_2)$ was tuned to the series resonant frequency of the ystal [the pole of $Y(p)$] in Fig. 6.7–2. We might ask what happens to the amplitude d frequency of the oscillations as the tuned circuit is detuned or mistuned.

On a small-signal basis, in the vicinity of the pole in $Y(p)$ of the crystal, $Y(j\omega)$ losely approximated by

$$Y(j\omega) \approx \frac{\omega_0^2 C}{2(j\delta\omega + \omega_0/2Q_L')},$$

ere

$$\delta\omega = \omega - \omega_0, \qquad \omega_0 + \omega \approx 2\omega_0, \qquad \text{and} \qquad Q_L' = \frac{\omega_0 L}{r + 1/g_{\text{in}Q}}.$$

normalized form of the variations in the magnitude and phase of this function given by the $Q = \infty$ curves of Figs. 2.2–7 and 2.2–8. The same normalized curves also valid for the impedance of the tuned circuit near ω:

$$Z(j\omega)_t = \frac{1}{2C_t(j\delta\omega + \omega_0/2Q_T)}, \qquad Q_T \approx \omega_0 R_L C_t, \tag{6.7–9}$$

ere the subscript t is used to differentiate the parameters of the tuned circuit from se of the crystal.

If we detune the tuned circuit by $\omega_0/8Q_T$ (2 kHz in the example where $Q_T = 100$ and $\omega_0 = 10^7$), we reduce the magnitude of $Z(j\omega)$ from R_L to $0.970R_L$ and introduce a phase shift of 14°. To cause a return to zero net phase shift, the operating frequency must shift until the crystal phase shift supplies the compensating 14°. Since the effective crystal Q is much higher than the tuned-circuit Q, the frequency need only shift by $\delta\omega = \omega_0/8Q_L'$ (16 Hz when $\omega_0 = 10^7$ and $Q_L' = 12,800$)†; however, this frequency shift causes an equivalent normalized decrease in the magnitude of $Y(j\omega)$. Thus, if we let

$$N = \frac{1}{\sqrt{1 + 4Q_T^2(\delta\omega_t/\omega_0)^2}}, \tag{6.7–10}$$

which is the normalized decrease in the magnitude of the tuned-circuit impedance as well as the decrease in the magnitude of the crystal admittance, then Eq. (6.7–7) becomes

$$\frac{G_{\text{in}}(y)}{g_{\text{in}Q}} = \left(\frac{N}{\alpha n N^2 R_L - r}\right)\frac{1}{g_{\text{in}Q}}. \tag{6.7–11}$$

If $\delta\omega_t$ is excessive, then $G_{\text{in}}(y)/g_{\text{in}Q}$ exceeds unity and oscillations cease (cf. Fig. 6.7–4). Thus detuning the tuned circuit causes only a small shift in frequency but may cause an appreciable change in amplitude. With $N = 0.97$,

$$\frac{G_{\text{in}}(y)}{g_{\text{in}Q}} = \frac{54}{59}, \qquad y \approx 0.58, \qquad \text{and} \qquad V_t = \alpha y I_{EQ} N R_L \approx 2.5\ \text{V};$$

with proper tuning $V_t \approx 3.7$ V. Thus good amplitude stability in this circuit demands reasonable frequency stability in the tuned circuit. If we assume that transistor reactances are negligible, then tuning the tuned circuit for maximum output amplitude yields operation at the series resonant frequency and thus minimizes the effects of mistuning on the amplitude.

Numerous variations of this basic circuit are apparent. For example, if we isolate the crystal from the emitter by an emitter follower and then use the same emitter resistor for both the oscillator transistor and the emitter follower, then a differential-pair oscillator results. Figure 6.7–5 illustrates such a circuit.

In such circuits it will often be true that R_B is small with respect to the "average" large-signal input impedance of Q_2. When this is the case, the driving voltage will still be sinusoidal and, except for the need to account for loss in gain through the R_B, r_{crystal} attenuator, the amplitude calculation will be identical to that of Section 6.5.

If R_B becomes too large then an exact calculation will require an expansion similar to Eq. (6.7–4) for the differential pair. While such a calculation is quite possible, it seldom appears to be necessary, and hence the results are not presented here.

---

† We assume here that $\delta\omega$ is negligible in comparison with $\delta\omega_t$, so that the effect of $\delta\omega$ on the magnitude and phase of the tuned circuit may be neglected.

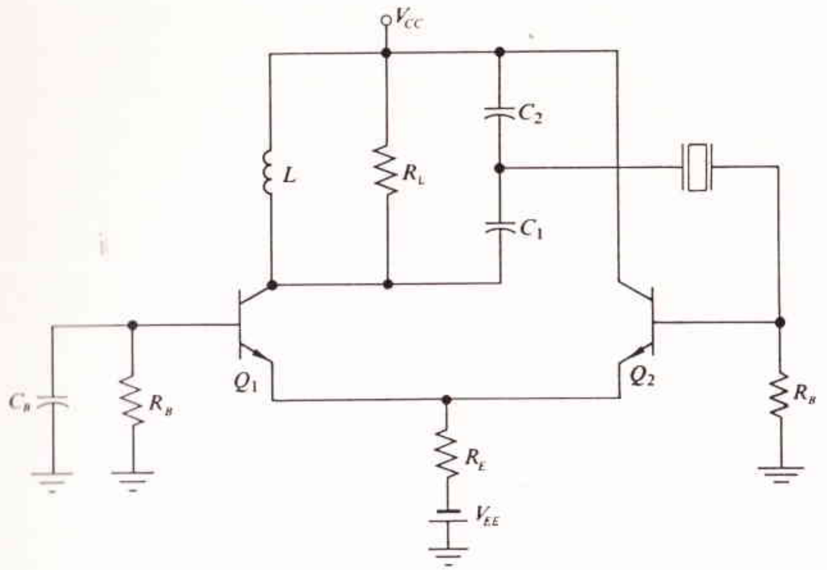


Fig. 6.7–5 Differential-pair crystal oscillator.

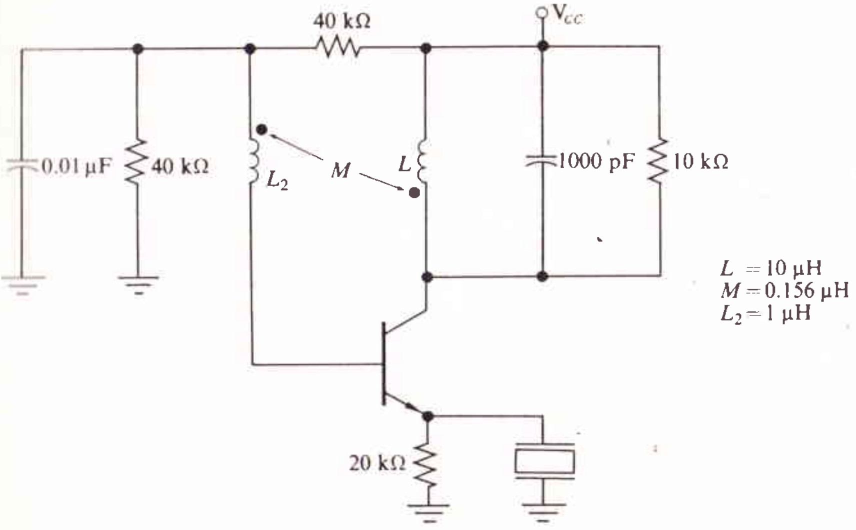


Fig. 6.7–6 Another series mode crystal oscillator.

If the crystal overtone frequencies are properly suppressed, then operation of this cuit is possible with commercial integrated-circuit differential pairs in which a d resistor replaces the tuned circuit and the combination of the crystal series sistor and the base input resistor replaces the capacitive attenuator.

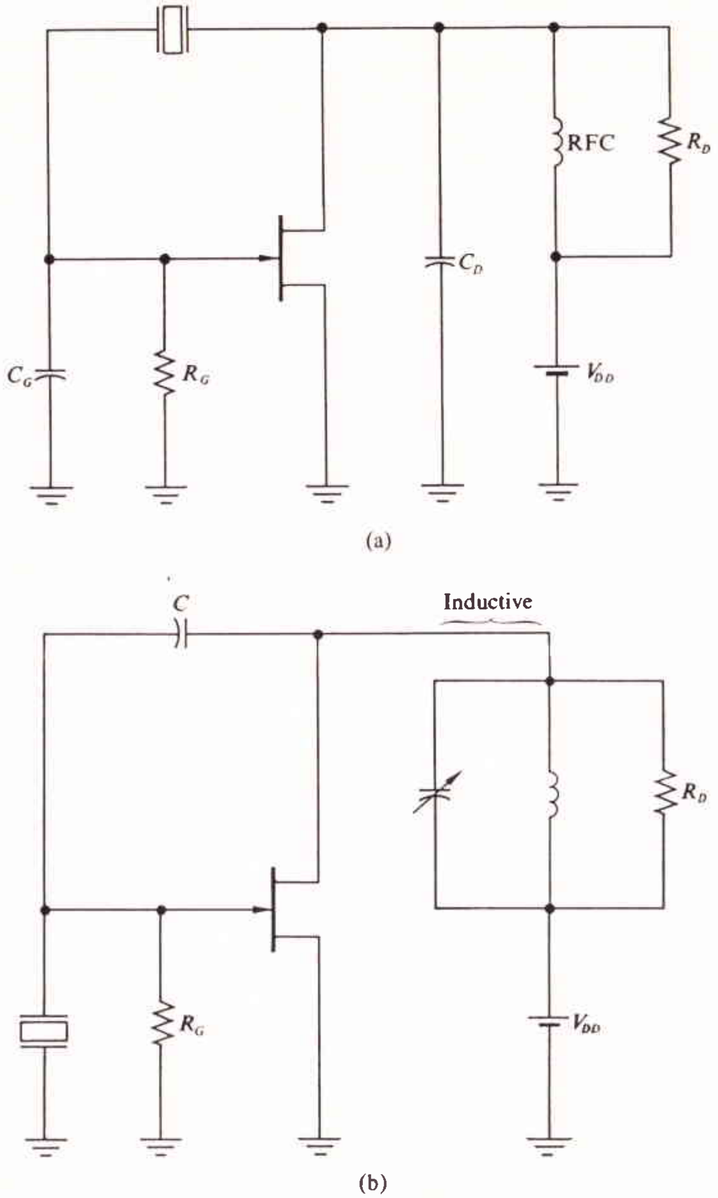


Fig. 6.7–7 (a) Pierce circuit (Colpitts-like). (b) Miller circuit (Hartley-like).

It is apparent that, for satisfactory operation in such a circuit, we require a much higher crystal series resistance than in the previous cases. This is true because the effective Q_L' is reduced to $r/(R_L + R_B + r)$ of its initial value, where r is the crystal series resistance and R_L and R_B are the circuit output and input resistances respecively. If $R_L = 4\,\text{k}\Omega$, $R_B = 500\,\Omega$, and $r = 10\,\text{k}\Omega$, then only a $\frac{2}{3}$ degradation of $_L$ results.

Another variation of the single-ended circuit is shown in Fig. 6.7–6. If $= M/L = 1/50$ and $r = 125\,\Omega$, then this circuit has the same output amplitude as he circuit of Fig. 6.7–3. The calculation of this amplitude is performed in exactly he fashion outlined previously. Intuitively we see that the circuit can have adequate oop gain to cause oscillations only when ω is near the series resonant frequency of he crystal (i.e., when the low series crystal resistance bypasses the large emitter mpedance).

Figure 6.7–7 illustrates two possible high-impedance mode FET (or vacuum-tube) rystal oscillator circuits. In both circuits the crystal behaves like an inductor and he operating frequency is near, but slightly below, the pole in $Z(p)$. For the Pierce ircuit,

$$G_m \approx \frac{C_D}{C_G}G_G + \frac{C_G}{C_D}G_D, \tag{6.7–12}$$

vhich for $C_D \approx C_G$ and $R_G \gg R_D$ leads to $G_m \approx 1/R_D$, where R_D is the load resistor. The required G_m is much smaller than for the series mode transistor; however, since he available G_m from an FET or vacuum tube is also much smaller, the combination s reasonable. That the operating frequency must lie slightly below the pole frequency seen by considering the crystal as an equivalent inductance which resonates with he series combination of C_D and C_G. Only within several multiples of ω_{pole}/Q_L elow ω_{pole} will the crystal have a sufficient inductive reactance.

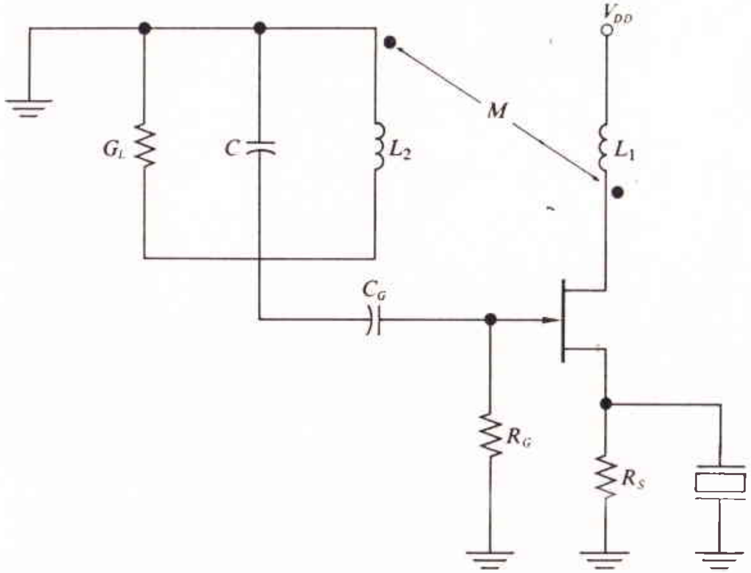


Fig. 6.7–8 Series mode FET crystal oscillator.

Similar arguments hold for the Miller circuit, where C resonates with an equivalent inductance from the crystal in series with an equivalent inductance from the drain-tuned circuit. The algebra is more complicated in this case and lends little insight into the circuit operation; hence it is not presented here.

Figure 6.7–8 shows a series mode crystal oscillator circuit which is a variation of the FET circuit of Fig. 6.6–1. In this case the small-signal g_m, with the crystal removed, is reduced to $g_m/(1 + g_mR_s)$. If R_s is sufficiently large so that Eq. (6.6–3) cannot be satisfied even for zero gate bias, then, with the crystal replaced, oscillations can occur only near the crystal's series resonant frequency.

As in the bipolar case, the source current is forced to be sinusoidal; hence again a complete calculation necessitates finding the fundamental source voltage as a function of the ac and dc currents.

6.8 SQUEGGING

In self-limiting sine-wave oscillators where the feedback loop is closed with a coupling capacitor, it is possible to have an interaction between the time constants of the bias and coupling circuits of the loop and the time constants of the high-frequency tuned circuits of the loop such that a self-produced amplitude modulation of the high-frequency oscillation occurs. Such self-modulated behavior is known as squegging. The low-frequency variations of the envelope of the high frequency may be sinusoidal or exponential.

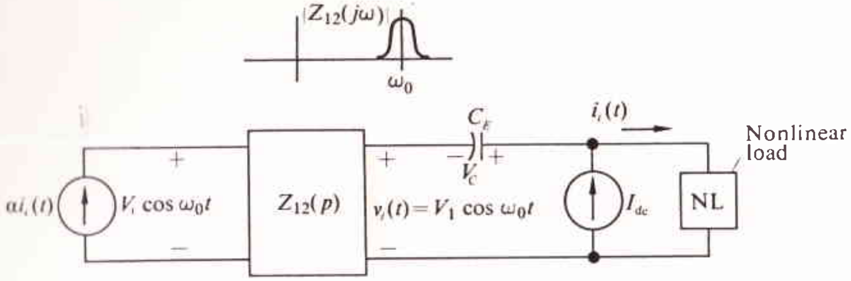
In this section we indicate an approach that either allows us to design for deliberate squegging or, as is more often the case, guarantees that it will not occur. First we shall outline a general method, and then we shall apply it to a specific transistor circuit to obtain definite numerical results.

In our analysis we assume that conditions exist for normal high-frequency oscillations to occur and that these oscillations have been set up. We then examine the small-signal low-pass transfer function that governs the amplitude of the envelope of these oscillations. If the low-pass transfer function has only left-half plane poles, then after an initial transient at $t = 0$, the amplitude remains constant at its desired value. However, if this low-pass transfer function has right-half plane complex conjugate poles, then sinusoidal oscillations of the amplitude occur; and if it has a right-half plane real-axis pole, then exponential buildups and decays of the amplitude occur.

Since we want some information about the onset of squegging, we are interested in the perturbation of the parameters relating to the envelope around their normal steady-state values. Consider first the generalized self-limiting capacitively coupled oscillator shown in Fig. 6.8–1, for which the self-limiting transistor oscillator shown in Fig. 6.8–2 is a special case. If for the circuit of Fig. 6.8–1 we assume a small perturbation $v_1(t)$ on the envelope of $v_i(t)$, that is,

$$v_i(t) = [V_1 + v_1(t)]\cos\omega_0 t,$$

then a small perturbation $i_1(t)$ results in the envelope of the fundamental component of $i_i(t)$. We note, however, that the right-half portion of the circuit of Fig. 6.8–1 is identical to the circuit of Fig. 5.6–1; thus $i_1(t)$ may be related to $v_1(t)$ by the admittance $Y_1(p)$ given by Eq. (5.6–3). In addition, if $Z_{12}(p)$ is a symmetric narrowband filter centered about ω_0 (which it almost always is), then the results of Section 3.3 tell us that $v_1(t)$ is related to the envelope variation of the fundamental component of the



$i_c(t) = I_0 + I_1 \cos \omega_0 t + I_2 \cos 2\omega_0 t + \cdots$

$A_L(j\omega_0) = 1$

Fig. 6.8–1 Generalized self-limiting capacitively coupled oscillator (nonlinear load).

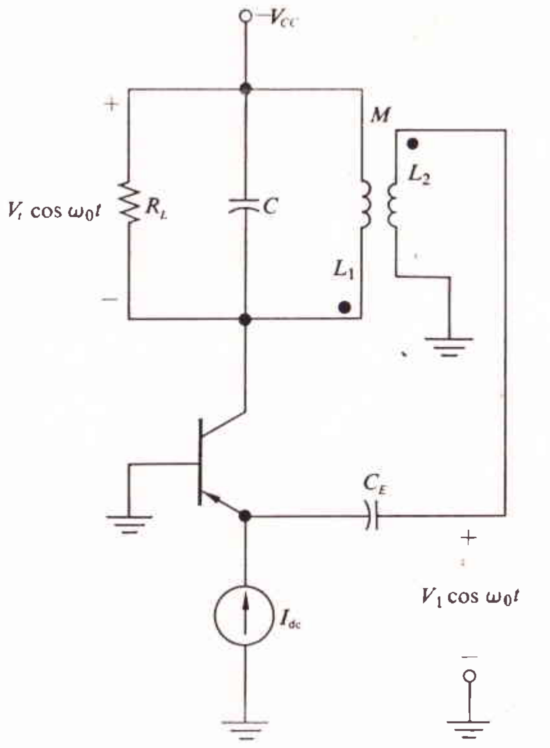
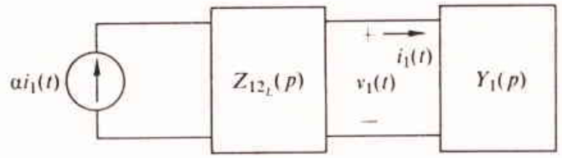


Fig. 6.8–2 Transistor oscillator. (It is assumed that collector saturation does not occur and that normal high-frequency oscillations are possible.)



$Z_{12_L}(p) =$ low-pass equivalent of $Z_{12}(p)$

$$Y_1(p) = \frac{G_{11}\left[p + \frac{1}{C}\left(G_{00} - \frac{G_{10}G_{01}}{G_{11}}\right)\right]}{p + \frac{G_{00}}{C}}$$

$$A_{L\,\text{env}}(p) = \alpha Y_1(p) Z_{12_L}(p)$$

Fig. 6.8–3 Envelope feedback circuit.

current driving the filter by the low-pass equivalent transfer function $Z_{12_L}(p)$ of the filter. The complete low-frequency feedback circuit relating $i_1(t)$ and $v_1(t)$ is shown in Fig. 6.8–3, from which it is apparent that

$$A_{L\,\text{env}}(p) = \alpha Y_1(p) Z_{12_L}(p). \tag{6.8–1}$$

If $1 - A_{L\,\text{env}}(p) = 0$ contains roots in the right-half complex plane, then squegging results. Even if no right-half plane roots exist, the positions of the left-half plane roots give us a good idea of the envelope transients when power is applied to the oscillator at $t = 0$.

To carry the analysis further, we require expressions for $Y_1(p)$ and $Z_{12_L}(p)$. For the transistor oscillator of Fig. 6.8–2 (or any of the self-limiting transistor oscillators discussed in Section 6.4), $Y_1(p)$ is given by (cf. Eq. 5.6–6)

$$Y_1(p) = \frac{2g_{\text{in}}\left[1 - \frac{I_1(x)}{xI_0(x)}\right]\left[p + \frac{\lambda(x)}{\tau}\right]}{p + \frac{1}{\tau}},$$

where $\tau = C_E/g_{\text{in}}$, $g_{\text{in}} = qI_{\text{dc}}/kT$, and $\lambda(x)$ is given by Eq. (5.6–7) and plotted vs. x in Fig. 5.6–2. In addition, if $Q_T = \omega_0 C R_T > 10$, then

$$Z_{12_L}(p) = \frac{\frac{nR_T}{\tau_L}}{p + \frac{1}{\tau_L}}, \tag{6.8–2}$$

where $\tau_L = 2R_T C$, R_T is the total loading across the tuned-collector circuit, and $n = M/L_1$. As is required of a low-pass equivalent, $Z_{12_L}(0) = Z_{12}(j\omega_0) = nR_T$.

Now with the specific values for $Z_{12_L}(p)$ and $Y_1(p)$, $A_{L\,\text{env}}(p)$ reduces to

$$A_{L\,\text{env}}(p) = \frac{\frac{2g_m nR_T}{\tau_L}\left[1 - \frac{I_1(x)}{xI_0(x)}\right]\left[p + \frac{\lambda(x)}{\tau}\right]}{\left(p + \frac{1}{\tau_L}\right)\left(p + \frac{1}{\tau}\right)}. \tag{6.8–3}$$

wever, since the circuit is assumed to be oscillating at ω_0, $A_L(j\omega_0) = 1$, or ivalently,

$$G_m(x)nR_T = g_m\frac{2I_1(x)}{xI_0(x)}nR_T = 1 \qquad \text{or} \qquad g_m nR_T = \frac{xI_0(x)}{2I_1(x)}.$$

th this additional restriction, $A_{L\,\text{env}}(p)$ may be expressed as

$$A_{L\,\text{env}}(p) = \frac{\dfrac{A(x)}{\tau_L}\left[p + \dfrac{\lambda(x)}{\tau}\right]}{\left(p + \dfrac{1}{\tau_L}\right)\left(p + \dfrac{1}{\tau}\right)}, \tag{6.8–4}$$

ere

$$A(x) = \frac{xI_0(x)}{I_1(x)} - 1.$$

pole-zero diagram of $A_{L\,\text{env}}(p)$ is presented in Fig. 6.8–4, in which the root locus of $- A_{L\,\text{env}}(p) = 0$ with increasing $A(x)$ is indicated. Clearly the possibility of squegging ists.

The roots of $1 - A_{L\,\text{env}}(p) = 0$ reach the imaginary axis when the coefficient of e p-term is zero. Therefore, to keep the roots in the left-half plane and thus to event squegging, we require

$$\frac{1}{\tau} + \frac{1}{\tau_L} \geq \frac{A(x)}{\tau_L} \tag{6.8–5a}$$

r

$$\frac{\tau_L}{\tau} = \frac{2R_TC}{C_E/g_{\text{in}}} = \frac{2Q_T}{\omega_0 C_E/g_{\text{in}}} \geq A(x) - 1 \tag{6.8–5b}$$

r

$$\frac{C}{C_E} \geq \frac{1}{2R_Tg_{\text{in}}}[A(x) - 1]. \tag{6.8–5c}$$

igure 6.8–5 shows a plot of $A(x) - 1 = xI_0(x)/I_1(x) - 2$.

For example, if an oscillator is designed to operate at a given x, with a given uned circuit capacitance C and load R_T, and at a given dc bias which fixes g_{in}, then there will be a maximum value of C_E at which squegging may be avoided.

From Fig. 6.8–5 we see that for small x, say less than 2, then $A(x) - 1$ is small and C_E may be relatively large without any danger of squegging. As x is increased, however, we may have difficulty in simultaneously satisfying Eq. (6.8–5) and the condition $\omega_0 C_E/g_{\text{in}} > 10$. (The latter condition is necessary if C_E is to be considered as an ac short circuit.)

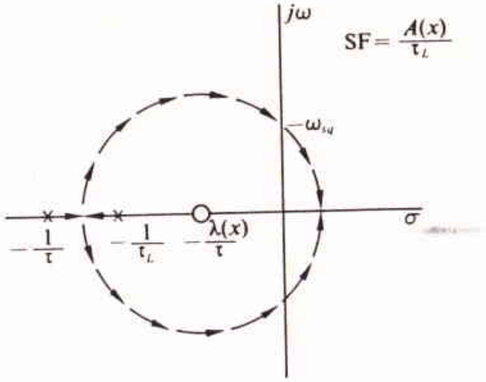


Fig. 6.8–4 Pole-zero diagram of $A_L(p)$ and root locus of $1 - A_{L\,\text{env}}(p) = 0$.

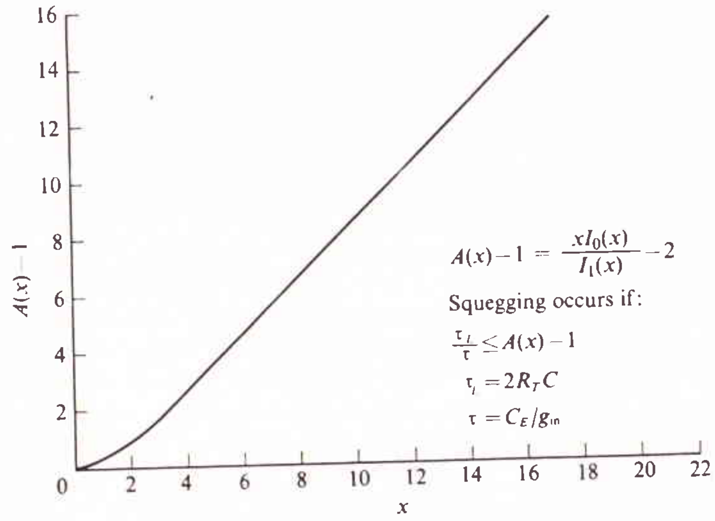


Fig. 6.8–5 Plot of $A(x) - 1$ vs. x.

For example, if $\omega_0 C_E/g_{\text{in}} = 20$ and $Q_T = 50$, then from Eq. (6.8–5b) and Fig. 6.8–5 we see that a design value of $x = 6.4$ will cause squegging. For proper non-squegging operation of such a circuit, either x must be kept small (below 6.4), or Q_T must be increased.

If squegging is allowed to occur, then its initial frequency of sinusoidal variation will be the frequency at which the roots of $1 - A_L(p) = 0$ cross the j-axis or

$$\omega_{\text{sq}} = \sqrt{\frac{1 - A(x)\lambda(x)}{\tau_L\tau}} = \frac{\omega_0\sqrt{1 - A(x)\lambda(x)}}{\sqrt{2Q_T\omega_0 C_E/g_{\text{in}}}}. \tag{6.8–6}$$

By combining terms we can write

$$1 - A(x)\lambda(x) = 2 + \frac{xI_1(x)}{I_0(x)} - \frac{xI_0(x)}{I_1(x)}.$$

Table 6.8–1 lists several values of $\sqrt{1 - A(x)\lambda(x)}$.

Table 6.8–1

x	$\sqrt{1 - A(x)\lambda(x)}$
0.0	0.00
1.0	0.46
2.0	0.73
3.0	0.86
6.0	0.95
10.0	0.97
∞	1.00

By using the large-x expansions for $I_1(x)$ and $I_0(x)$, one can show that, to within 5% for $x \geq 3$,

$$\sqrt{1 - A(x)\lambda(x)} \approx 1 - \frac{5}{16x}.$$

Therefore, since the exact frequency of squegging is seldom of great importance and since squegging is most likely for $x \geq 3$, it seems reasonable to rewrite Eq. (6.8–6) as

$$\omega_{sq} \approx \frac{\omega_0(1 - 5/16x)}{\sqrt{2Q_T(\omega_0 C_E/g_{in})}}. \qquad (6.8\text{–}7)$$

For the previous example, where $Q_T = 50$, $\omega_0 C_E/g_{in} = 20$, and $x = 6.4$, we find that

$$\omega_{sq} \approx \omega_0/47.5.$$

In actual circuits the values to be used for C and for C_E are usually obvious from inspection. For example, in the Colpitts-type circuit $C = C_1C_2/(C_1 + C_2)$ and $C_E = C_1 + C_2$. If a bypass capacitor C_B is placed from the base to ground (as shown in Fig. 5.4–1), then $C_B/(1 - \alpha)$ must be added in series with C_E to determine the total coupling capacitor.

A similar technique, starting with the results of Fig. 6.8–3 and evaluating the four partial derivatives, either analytically or experimentally, will lead to a similar set of results for vacuum-tube† or FET oscillators. In all cases squegging can be stopped, if it occurs, by some combination of

a) reducing the size of the coupling capacitor (if other restrictions will allow this reduction),
b) raising the tuned-circuit Q, and
c) reducing the active-element driving voltage.

---

† Two sources that consider squegging in the vacuum tube case in some detail are: (1) W. R. MacLean, "Criteria for the Amplitude Stability of a Power Oscillator," *Proc. IRE*, **42**, pp. 1784–1791 (Dec. 1954); and (2) B. G. Dammers, J. Haantjej, J. Otte, and H. Van Suchtelen, "Squegging Oscillators," in *Application of the Electronic Valve in Radio Receivers and Amplifiers*, Book IV. N. U. Philip's Gloeilampenfabrieken, Eindhoven, Netherlands (1950), Chapter II E, pp. 227–250. In the vacuum tube case, it is not possible to arrive at nearly such a compact nor universal result as we have obtained here for the transistor.

6.9 BRIDGE OSCILLATORS

Bridge oscillators utilize the transfer function of a bridge network to produce a pair of high-Q right-half plane complex zeros. As was pointed out in Section 6.3, such a set of zeros contributes directly to the indirect frequency stability of the circuit.

In this section two basic bridge circuits are discussed. The first is the fixed-frequency Meacham bridge, in which the already high Q of a crystal resonator is multiplied by several additional orders of magnitude to yield a *very* stable circuit. The second is the wide-range variable-frequency circuit known as the Wien bridge oscillator.† In the Wien bridge oscillator the aim is not to achieve extreme frequency stability, but rather to achieve the stability and waveform purity of an LC circuit of reasonable Q while retaining the wideband tuning capability of RC networks.‡

A bridge oscillator consists of three main elements:

a) the bridge circuit,
b) a linear amplifier, and
c) a bridge-balancing circuit or element.

As we shall see, the desirable properties of the bridge oscillators are obtained only if the bridge is operated as close to balance as is possible. Since the amplifier gain times the bridge error signal must be a constant if stable operation is to be achieved, and since the amplifier gain A cannot be assumed to be a constant, we achieve amplitude control by sampling the output amplitude and using this signal to provide exactly the necessary bridge unbalance.

Figure 6.9–1 illustrates both bridge circuits under discussion. With vacuum-tube circuitry the self-balancing effect is often achieved with a positive-temperature-coefficient resistor (a light bulb) in the upper half of the negative feedback arm (element R_3). In addition to being somewhat sensitive to ambient temperature, such elements often require more power for their operation than the total power available from a small semiconductor operational amplifier; hence in semiconductor circuitry it is usually more reasonable to obtain the self-balancing effect with an FET, a light-sensitive resistor, or perhaps a diode acting as a controlled resistor.

While we could operate a bridge oscillator with a simple differential-pair circuit as an amplifier, the desirable properties of the circuit require high voltage gains;

---

† The literature is full of RC feedback oscillators that are misnamed Wien bridge circuits; these circuits do not contain any bridge structure and hence do not have the properties described here.

‡ Wien bridge oscillators are common with 10/1 variations in frequency per dial turn and overall coverage of $10^6/1$, say from 2 Hz to 2 MHz. With an RC network a 10/1 variation of R or C gives a 10/1 variation of frequency, while in the LC case it yields only a 3.16/1 variation. In addition, for frequencies below several kilohertz, linear inductors become both unwieldy and expensive.

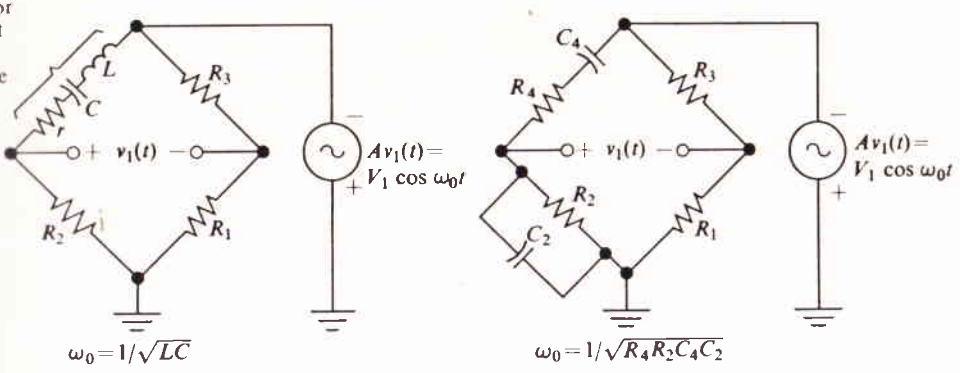


Fig. 6.9–1 (a) Meacham bridge. (b) Wien bridge.

nce we assume an ideal operational amplifier with a sufficiently high-impedance fferential input that bridge loading may be neglected and with a sufficiently low tput impedance that the bridge does not load the amplifier. So long as we restrict e output signal level to several volts and the frequency range to below, say, 10 MHz, en such amplifiers are widely available at relatively low prices. In practical circuits, rticularly at low frequencies where impedance levels are high, one often uses an ET pair at the input to the differential amplifier to further reduce the effect of ading.

leacham Bridge

there is zero phase shift through the amplifier, then the Meacham bridge oscillator Fig. 6.9–1a operates extremely close to the series resonant frequency of the crystal, d thus $A_L(j\omega_0)$ is given by

$$A_L(j\omega_0) = +A\left(\frac{R_1}{R_1 + R_3} - \frac{R_2}{R_2 + r}\right). \tag{6.9–1}$$

We define $N = R_2/r$, $M = R_3/r$, and $\delta = NM - R_1/r$. If $(N + 1)M \gg |\delta|$, then able oscillations occur $[A_L(j\omega_0) = 1]$ when

$$\delta = -\frac{(N + 1)^2 M}{A}. \tag{6.9–2}$$

Note: $\delta = 0$ corresponds to perfect bridge balance, while a slightly negative value δ is obtained when $v_1(t)$ is a small negative voltage relative to its defined polarity.]

For example, if the bridge arms are nearly equal, so that $M = N = 1$, then $= -4/A$; hence if $A > 100$, certainly $|\delta| \ll (N + 1)M$ and our assumption is stified. It will be the job of our bridge-balancing circuitry to adjust δ so that q. (6.9–2) is always satisfied regardless of variations in A.

The complete expression for the loop gain of the Meacham bridge is given by

$$A_L(p) = +\frac{R_1 A}{R_1 + R_3}\,\frac{p^2 - 2\alpha_1 p + \omega_0'^2}{p^2 + 2\alpha_2 p + \omega_0'^2}, \tag{6.9–3}$$

where

$$\omega_0' = \frac{1}{\sqrt{LC}} = \omega_0, \qquad \alpha_1 = \frac{1}{2L}\left(\frac{R_2 R_3}{R_1} - r\right), \qquad \text{and} \qquad \alpha_2 = \frac{1}{2L}(r + R_2).$$

The pole-zero diagram of $A_L(p)$ is given in Fig. 6.1–7(b). [The reader should convince himself that the equation for $A_{\min}$ in the figure is identical to Eq. (6.9–2).] It is apparent that the effective Q of the zeros of $A_L(p)$ has the form

$$Q_z = \frac{\omega_0'}{2\alpha_1},$$

while the Q of the poles is given by

$$Q_p = \frac{\omega_0'}{2\alpha_2}.$$

If we define the Q of the crystals as $Q_L = \omega_0' L/r$, then we obtain the following interesting ratios:

a) $$\frac{Q_z}{Q_p} = \frac{\alpha_2}{\alpha_1} = \frac{R_1(R_2 + r)}{R_2 R_3 - rR_1} = \frac{(NM - \delta)(N + 1)}{\delta} \approx \frac{MN(N + 1)}{\delta}$$

$$= \frac{NA}{N + 1} \qquad \text{for } A > 100,$$

which is completely independent of the purely resistive side of the bridge, and

b) $$\frac{Q_z}{Q_L} = \frac{R_1 r}{R_2 R_3 - rR_1} = \frac{NM - \delta}{\delta} \approx \frac{NM}{\delta} = \frac{NA}{(N + 1)^2}. \tag{6.9–4}$$

Clearly, for a given crystal, the greater the ratio Q_z/Q_L, the greater the indirect frequency stability (see Eq. 6.3–15).

The ratio Q_z/Q_L is a function of N which is geometrically symmetrical about $N = 1$, where it obtains its maximum value of $A/4$. By making $N = 4$ or $\frac{1}{4}$, we reduce Q_z/Q_L to 64% of its maximum value; by making $N = 10$ or $\frac{1}{10}$, we reduce the ratio to 33% of its maximum value.

If we wish to vary N from unity, we suspect that it would be desirable to increase it, since this would raise the bridge impedance level; a higher impedance level would in turn make the bridge easier to drive, reduce the crystal current, and reduce the crystal dissipation.

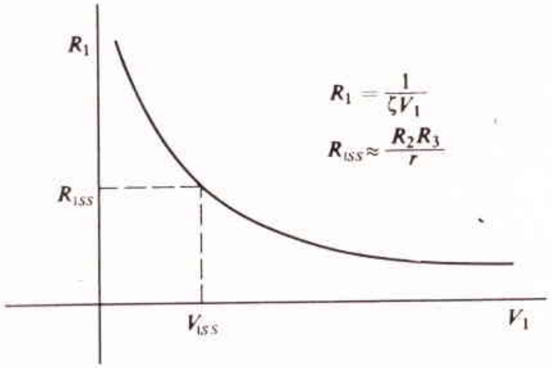
From Eq. (6.9–4) we see that for $N = 1$ a 60 dB amplifier gain can lead to a frequency stability enhancement 250 times that obtainable with the crystal alone. Of course, to benefit from this increased indirect frequency stability we must ensure ı degree of direct frequency stability; i.e., we must make sure that ω_0' of the l remains stable. This is usually accomplished by placing the crystal in an which maintains its temperature constant within a few degrees Celsius.

beautiful feature of the Meacham and Wien bridge oscillators is that the same s which lead to a large indirect frequency stability factor also cause the output tion amplitude V_1 to be insensitive to variations in the amplifier gain A; that is, ause the amplitude sensitivity function S_A to be small (cf. Section 6.4). Specifi-

$$S_A = \frac{(N+1)^2}{NA}; \qquad (6.9\text{–}5)$$

S_A is minimized with $N = 1$ and A as large as possible.

o demonstrate the validity of Eq. (6.9–5), we assume that R_1 is the bridge-cing element which decreases with increasing output voltage amplitude V_1, own in Fig. 6.9–2. Although a hyperbolic relationship is shown between V_1 and e same results may be obtained for any similar relationship between V_1 and R_1. ever, it is apparent that such a voltage resistance relationship does lead to tude stability. If V_1 should decrease for some reason, R_1 increases and in turn s $A_L(j\omega_0)$ [given by Eq. 6.9–1)] to increase, thus restoring the output oscillation tude.



.9–2 Relationship between the bridge balancing resistor R_1 and the output voltage ampli-

he amplitude V_{1SS} at which V_1 stabilizes may be found directly from the curve . 6.9–2. If we assume that A is large (which it must be for a well-designed circuit), δ is small, since Eq. (6.9–2) must be satisfied in the steady state. A small value implies bridge balance or, equivalently, $R_{1SS} \approx R_2R_3/r$; hence the value of V_1 ı results in $R_1 \approx R_2R_3/r$ on the curve of Fig. 6.9–2 is V_{1SS}. Clearly, if we desire cific output amplitude, we must choose our voltage-controlled resistance to be l to R_2R_3/r for the desired output amplitude. Or, equivalently, for a given characteristic, R_3 must be chosen such that $R_3 = R_1/N$ for the desired output

amplitude. For the optimum case of $N = 1$, $R_3 \approx R_{1SS}$. From the above reasoning it is apparent that, if A is sufficiently large to keep the bridge balanced and R_2, R_3, r, and the R_1-V_1 characteristic are stable, then the output amplitude is quite stable.

We now evaluate S_A quantitatively by noting that

$$S_A \approx \frac{\partial V_1}{\partial A}\frac{A}{V_1} = \frac{A}{V_1}\frac{\partial V_1}{\partial R_1}\frac{\partial R_1}{\partial \delta}\frac{\partial \delta}{\partial A}. \qquad (6.9\text{–}6)$$

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION
Y BIBLIOTECA
MONTEVIDEO - URUGUAY

Since $V_1 = 1/\zeta R_1$,

$$\frac{\partial V_1}{\partial R_1} = -\frac{1}{\zeta R_1^2}.$$

In addition since $R_1 = r(NM - \delta)$,

$$\frac{\partial R_1}{\partial \delta} = -r;$$

and, from Eq. (6.9–2),

$$\frac{\partial \delta}{\partial A} = \frac{(N+1)^2 M}{A^2}.$$

Consequently,

$$S_A = \frac{(N+1)^2 Mr}{AR_1} \approx \frac{(N+1)^2}{AN} \qquad \text{(for large } A\text{)}.$$

Since ζ drops out of the expression for S_A, we see that the exact form of the V_1-R_1 characteristic is not critical so long as R_{1SS} produces the desired V_{1SS}.

Two practical Meacham bridge oscillator circuits in which R_1 is controlled inversely with the output amplitude are shown in Figs. 6.9–3 and 6.9–4. In Fig. 6.9–2 the output voltage drives a small light source which illuminates a light-sensitive

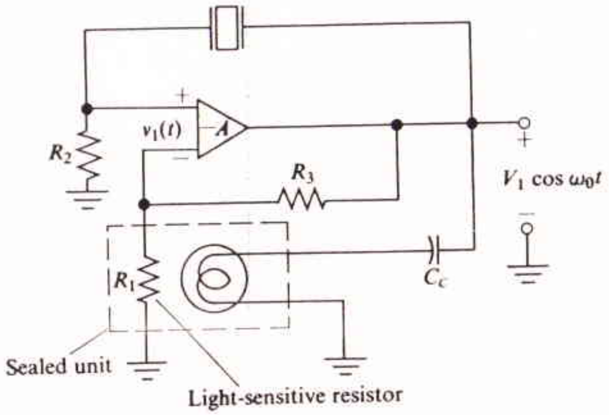


Fig. 6.9–3 Meacham bridge controlled by light-sensitive resistor.

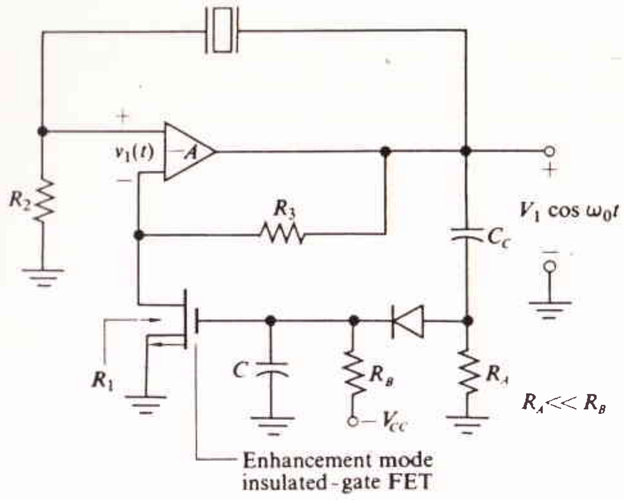


Fig. 6.9–4 Meacham bridge controlled by enhancement mode insulated-gate FET.

semiconductor resistor. As the output voltage and, in turn, the power to the light source increase, the semiconductor resistance decreases. Many sealed units containing both the light source and the semiconductor resistor are available commercially with nominal resistance values (with no illumination) from several hundreds of ohms to several hundreds of kilohms. The Raytheon Corporation produces a line of such elements under the trade name of Raysistors.[†]

In Fig. 6.9–4 the output voltage drives a peak envelope detector of the form discussed in Section 10.5 to develop a dc voltage proportional to V_1. This dc voltage in turn inversely controls the drain-source resistance of the enhancement mode insulated-gate FET. Since the drain-source resistance in this circuit remains linear only for drain-source voltages under 200 mV, there will be some distortion in the voltage across the FET if V_1 is designed to stabilize at too high a level. Fortunately, however, the Meacham and Wien bridge circuits possess a large amount of negative feedback (as the interested reader may calculate) at the harmonics of ω_0; therefore, very little of the harmonic components appearing across the FET appear at the output, especially if A is large. Thus, for large A, nonlinearities introduced by R_1 do not appreciably destroy the sinusoidal purity of the output.

Although maximum indirect frequency stability and minimum amplitude sensitivity both demand that $N = R_2/r = 1$, the permissible crystal dissipation sometimes requires us to select an N different from 1 and make up the deficiency in S_F and S_A with increased amplifier gain.

Excessive crystal current may cause the crystal to fracture. Flat statements about the size of a destructive current are difficult to make, since it may exceed 25 mA for a physically robust low-frequency unit and it may be less than 1 mA for a fragile high-frequency unit. Even without approaching the level of destruction, one normally wishes to keep the crystal dissipation below one milliwatt so as to minimize internal heating. Such heating usually causes shifts in frequency of the crystal resonances. The amount and directions of such shifts will be a function of the "cut" of the crystal.

For operation at series resonance, the crystal power is $V_1^2/2r(1 + N)^2$; hence, for $r = 100\ \Omega$ and a maximum allowable crystal power of 1 mW, we have

$$\frac{V_1}{1 + N} \leq 0.445 \qquad (6.9\text{–}7)$$

or

V_1, V	$N_{\min}$
1	1.25
2	3.5
4	8.0

Smaller values of r will place even more severe limitations on N.

Wien Bridge

For the Wien bridge circuit of Fig. 6.9–1(b) the loop gain is given by

$$A_L(p) = \frac{R_1 A}{R_1 + R_3} \frac{p^2 + p\left(\dfrac{1}{R_4 C_4} + \dfrac{1}{R_2 C_2} - \dfrac{R_3}{R_1 R_4 C_2}\right) + \dfrac{1}{R_4 R_2 C_4 C_2}}{p^2 + p\left(\dfrac{1}{R_4 C_4} + \dfrac{1}{R_2 C_2} + \dfrac{1}{R_4 C_2}\right) + \dfrac{1}{R_4 R_2 C_4 C_2}}. \qquad (6.9\text{–}8)$$

The pole-zero diagram of $A_L(p)$ is shown in Fig. 6.1–7(a).

Though there may be cases where the RC arms should not be identical, it is usually desirable to impose this restriction. With the special case where $R_2C_2 = R_4C_4 = R_4C_2 = 1/\omega_0'$ and the definition of R_3 as $R_3 = (2 + \delta)R_1$, Eq. (6.9–8) reduces to

$$A_L(p) = \frac{A}{3 + \delta} \frac{p^2 - p\delta\omega_0' + \omega_0'^2}{p^2 + 3p\omega_0' + \omega_0'^2}. \qquad (6.9\text{–}9)$$

This oscillator has j-axis poles at $\omega_0 = \omega_0'$ when

$$A = \frac{9}{\delta} + 3 \qquad \text{or} \qquad \delta = \frac{9}{A - 3}. \qquad (6.9\text{–}10)$$

This value for δ leads to a "que" for the complex zeros of

$$Q_z = \frac{1}{\delta} = \frac{A - 3}{9} \approx \frac{A}{9} \qquad (6.9\text{–}11)$$

when $A \gg 3$; hence, to achieve $S_F \approx 2Q_z = 100$, we must have $A = 450$.

[†] A possible disadvantage of the Raysistor as a control element is that it requires a relatively large amount of input current and also tends to be somewhat expensive.

In this case the limits on the allowable impedance levels for the bridge are set by maximum desirable capacitance and the problem of the amplifier input impedance ding the bridge. These problems become extreme at low frequencies; at high quencies, amplifier output impedance and input shunt capacitance may be limiting tors.

In general, the bridge input impedance can be kept to the order of several ndreds of ohms; hence the ordinary operational amplifier should drive it without liculty. Any of the control circuits discussed for the Meacham bridge case is an propriate replacement for R_1 in the Wien bridge.

0 THE ONE-PORT APPROACH TO OSCILLATORS

this section we examine another approach to the analysis and design of sine-wave cillators. Consider the four circuits in Fig. 6.10–1. As we shall see, all of them have potential to produce nearly sine-wave oscillations. They are representative of a iss of circuits—or of a way of looking at circuits—in which the amplifying device d its nonlinearition are lumped into a single, reactance-free, controlled source.

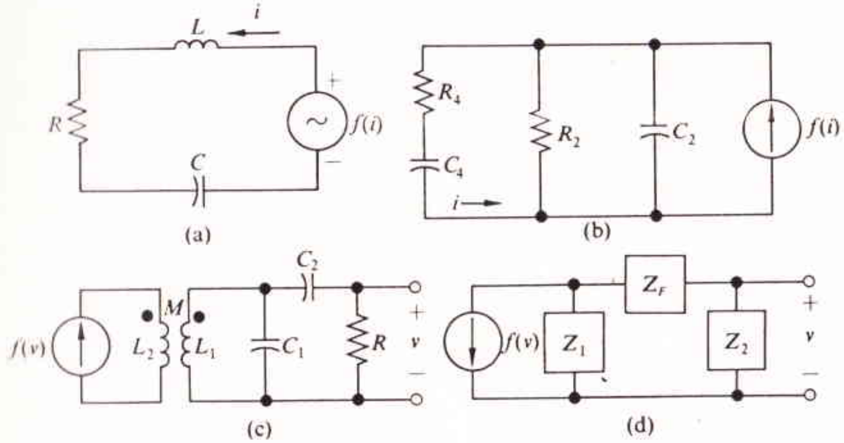


Fig. 6.10–1 Four "one-port" oscillators.

All of these circuits contain a filter, an ac energy source, and some form of positive edback; hence, if the loop gain is high enough, all of them are potential oscillators.

In this section we are interested not in proving that oscillation is possible, or in etermining the exact frequency or waveshape of oscillation, but in seeing how ossible nonlinearities in the controlled source may limit the amplitude of the cillations.

For the purposes of our basic explanation it is immaterial whether the nonlinear evice is constructed of separate elements such as FET's or transistors or whether is a single one-port device such as a tunnel diode. For example, the voltage-ontrolled current source in Fig. 6.10–1(d) might be a single bipolar or FET transistor,

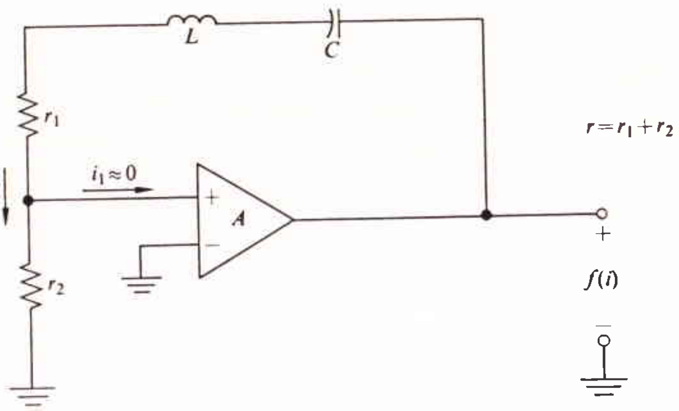


Fig. 6.10–2 Possible implementation of Fig. 6.10–1 (a).

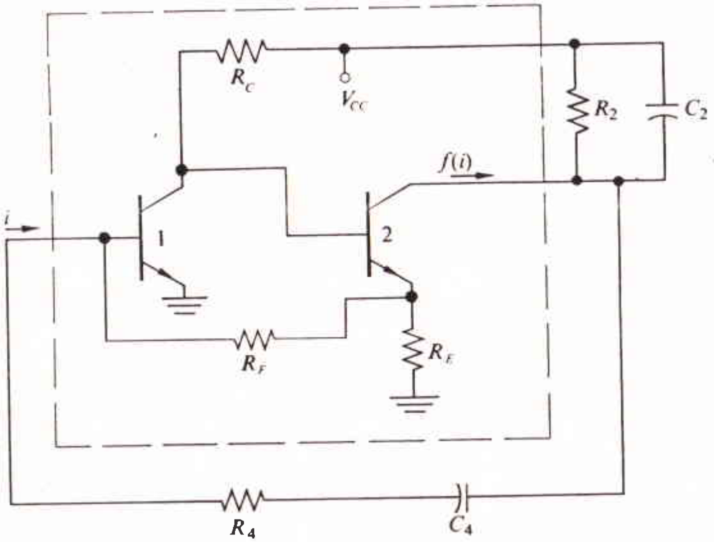


Fig. 6.10–3 Possible construction for Fig. 6.10–1 (b).

while the current-controlled voltage source in Fig. 6.10–1(a) might be constructed of a current-sensing resistor and an operational amplifier as illustrated in Fig. 6.10–2.

Figure 6.10–3 shows one of many ways to construct a circuit of the type shown in Fig. 6.10–1(b).

It is not our purpose at this point to develop the nonlinear characteristics of the various active "devices," but rather to calculate the oscillator performance given a calculated or measured "device" characteristic.

As a first example consider the circuit of Fig. 6.10–2, where the combination of operational amplifier and r_2 is assumed to have a characteristic of the form shown in Fig. 6.10–4.

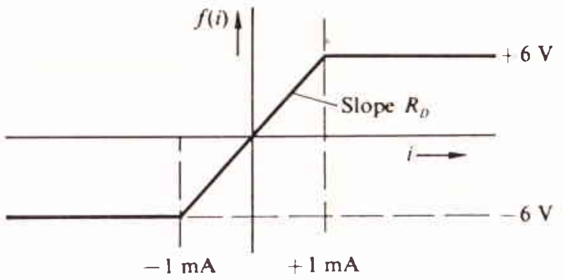


Fig. 6.10–4 Assumed characteristic for device in Fig. 6.10–1 (a) or Fig. 6.10–2.

First consider the small-signal situation, where the controlled source always operates in its linear region, $|i| \leq 1$ mA. In this region $f(i)$ may be replaced by a linear voltage source, $R_D i$.† With this source the loop gain $A_L(p)$ becomes

$$\frac{R_D pC}{p^2LC + prC + 1} = A_L(p),$$

where $r = r_1 + r_2$.

When $R_D = r$, the circuit has j-axis poles at $\omega_0 = \pm 1/\sqrt{LC}$; hence sinusoidal oscillations build up at this frequency. When $R_D < r$, the poles are in the left-half plane and, while transients may cause ringing, free oscillations do not occur. For $R_D > r$ the poles are in the right-half plane and the oscillations grow. In fact, in Fig. 6.10–1(a), if we merely replace the generator $f(i)$ by the negative resistor $-R_D$ (which is permissible, since the voltage source is linearly related to the current through it), we can perceive the previous results in an even more straightforward manner. When the net circuit series resistance is positive, the oscillations decay; when this resistance is negative, they grow. (Since for $|i| \leq 1$ mA any increase in i causes an increase in $f(i)$, the equivalent resistance $-R_D$ must indeed be negative.)

If initially $R_D = Ar_2 > r_1 + r_2$, then certainly oscillations build up until eventually at least the peaks of i exceed the linear region in Fig. 6.10–4. For the given characteristic, when $|i| > 1$ mA the $f(i)$ generator ceases to be a negative resistance and becomes a 6 V battery whose polarity depends on i. This action causes limiting in the buildup of i, and eventually an equilibrium condition is reached. It is this equilibrium state that we shall now find.

In general, when operation of the circuit shown in Fig. 6.10–2 extends beyond the linear region, $f(i)$ becomes a distorted sinusoidal voltage which may be written in the form

$$f(i) = V_0 + V_1 \cos \omega_0 t + V_2 \cos \omega_0 t + \cdots +.$$

However, if the Q_T of the tuned circuit is high enough that even the distorted voltage $f(i)$ still produces an almost sinusoidal current through r, then we may assume sinusoidal input currents of the form $I_1 \cos \omega_0 t$ for several values for I_1 and utilize Fig. 4.2–3 to calculate the harmonics in the subtracted sine-wave tips in order to calculate the output fundamental voltage V_1. Figure 6.10–5 illustrates the technique.

---

† In the circuit of Fig. 6.10–2, $R_D = Ar_2$.

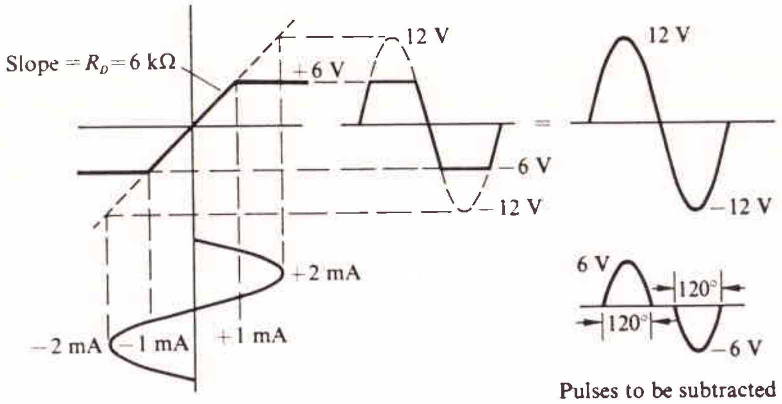


Fig. 6.10–5 Sinusoidal current drive.

The fundamental output voltage equals the extrapolated linear transfer value (neglecting the break points) minus the fundamental components in the sine-wave tip pulses that must be removed from the extrapolated value to obtain the actual characteristic. For the case shown,

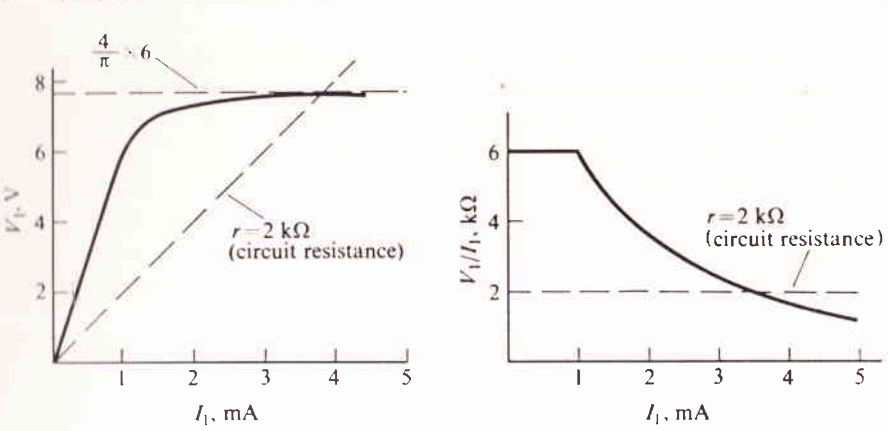
$$V_1 = [12 - 6 \times 0.39 - 6 \times 0.39] \text{ V} = 7.32 \text{ V},$$

where 0.39 is the function of 2ϕ read from the $n = 1$ curve of Fig. 4.2–3 for $2\phi = 120°$. By applying this technique several times we obtain a plot of V_1 vs. I_1 for the device. Now since the passive circuit also imposes a relationship between V_1 and I_1 of the form $I_1 = V_1/r$, we see that, as in previous sections, the operating amplitude is fixed by the simultaneous satisfaction of the active device and of the passive circuit requirements. Figure 6.10–6 illustrates several ways to present the results of this calculation.

The $(4/\pi) \times 6$ limit of V_1 in Fig. 6.10–6(a) is the limit of the fundamental value of a square wave (see Eq. 4.2–2). With $r = 2$ kΩ, this circuit will stabilize at about 3.8 mA $= I_1$, for which $V_1 = I_1 r = 7.6$ V. Now for this amplitude we can return to Fig. 4.2–3 and calculate the amplitudes of the first several harmonics in order to see how high Q must be for the assumption of a sinusoidal current to be valid in the first place. In this symmetric case, the second-harmonic voltage is zero, while the third-harmonic voltage is 2.4 V; hence, to reduce the third-harmonic current to 1% of the fundamental current, we must have

$$\frac{2.4}{7.6} \cdot \frac{n}{(n^2 - 1)Q_T} = 0.01 \qquad \text{or} \qquad Q_T \geq 12.$$

As r is increased toward 6 kΩ the overdrive becomes smaller and the Q requirement relaxes. If $r > 6$ kΩ, then no intersection occurs in Fig. 6.10–6 and operation ceases. In this case the net circuit resistance is always positive.



(a) (b)

Fig. 6.10–6 (a) V_1 vs. I_1 (from Fig. 6.10–5). (b) V_1/I_1 (from Fig. 6.10–5).

As we pointed out previously, Fig. 6.10–6(b) allows us to visualize the nonlinear controlled source as a negative resistor [remember that i flows up through $f(i)$ hence the generator is supplying rather than consuming ac power]. From this viewpoint the circuit will stabilize when the passive circuit resistance just balances the "average" negative resistance $-\bar{R}_D = V_1/I_1$ from the active circuit. Note that since all "negative-resistance" oscillators may be approached from the controlled-source viewpoint, there is no need to isolate them as a special case.

In this approach the feedback loop employed previously does not always appear explicitly; hence the previously employed criteria for oscillation (the loop gain being one) have been replaced by the criterion of zero net resistance [or j = axis poles—which is, of course, exactly what was achieved when $A_L(j\omega_0)$ reached unity].

A number of questions may arise at this point. What does one do when Q is not high? What sort of transients occur during the buildup of oscillations? What is the actual frequency of oscillation? (It will not be $1/\sqrt{LC}$ in Fig. 6.10–2, for example.) How does one deal with the case of smoothly varying device characteristics?

When the Q is not high enough to remove essentially all the harmonics, two approaches are possible. One is a perturbation technique whereby we initially assume a sine-wave drive and then find an approximate level of operation. Then, near that level, we determine the first several harmonics assuming a sine-wave drive, and after filtering these harmonics through the circuit we use the approximation to the actual output waveshape as a new driving function. With piecewise-linear characteristics this approach is not appealing.

For the smoothly varying case, we may approximate the function over what is estimated to be the useful region by as low an order of polynomial as appears plausible, and proceed. The iterative technique is more reasonable in this case.

Another approach which attacks the low-Q problem directly and which also yields results for transient buildups is the phase plane method described in Section 6.11. Section 6.12 then considers the effect of harmonic components on the operating frequency and thus relates waveshape distortion to variations in operating frequency.

6.11 THE PHASE PLANE APPROACH

Though the literature is full of presentations of the "phase plane" method, most of these general approaches turn out not to be easily applicable to practical oscillator problems of the type presented in Section 6.10.

This section presents a modified version of the Liénard method that leads to rapid solution of this type of problem. The specific method presented here is applicable to all second-order systems with nonlinear damping, that is, to systems in which the small-signal transfer function from the controlled source to the controlling variable has a single zero at the origin and only two poles (real or complex). Though the method as presented will not work directly for such circuits as the Meacham bridge, it is directly applicable to all the circuits of the previous section.

Chapter 2 of Hayashi† outlines more sophisticated graphical methods that are applicable to more complicated circuits.

The method presented here provides a graphical solution for the initial transient buildup and for the final operating waveshape of the oscillator. Its use implies that the reactances of the controlled-source (if any) have been separated out and combined with the circuit reactances. It also assumes that the circuit's dc Q-point is known and that it does not shift appreciably between the no-oscillation and full-oscillation cases. Once the output waveshape has been determined, it will be clear whether an average value of current or voltage does exist. If so, and if such an average value will have caused a Q-point shift (this will be a function of the bias circuitry), then the result is only an approximation. In such cases it is possible to pursue a perturbation approach through several steps toward a better estimate of the result. This is seldom undertaken, since the main purpose of the phase plane method is to provide rapid insight into the type of distortion likely to occur in oscillator circuits with "low-Q" filters.

The initial estimate, even if not exact, is usually sufficient for design purposes. The method handles both piecewise-linear and gradual nonlinearities with equal ease. In circuits with highly selective filters, however, one of the analytic methods described previously will give both quicker and more accurate results.

The result to be achieved is a graphical construction of a contour in a "phase plane" that will indicate whether oscillations are possible (a small-signal calculation will, of course, indicate this fact also) and, if so, how they build up from any particular set of initial conditions. A further construction derived from the first one gives a reasonable estimate of the circuit's waveshape and fundamental frequency.

In order to achieve these aims, a certain set of algebraic and trigonometric manipulations must be presented. The new reader would probably do well to skim through these once to see the conclusion and then return to unravel the details.

Basically what we are trying to do is obtain a solution of a nonlinear differential equation of the form

$$\frac{d^2x(t)}{dt^2} + a\frac{dx(t)}{dt} + bx(t) = c\frac{d}{dt}f(x), \tag{6.11-1}$$

---

† Chihiro Hayashi, *Nonlinear Oscillations in Physical Systems*. McGraw-Hill, New York (1964).

re x may be a voltage or a current and $f(x)$ is some known single-valued function The method involves the manipulation of the equation so as to remove time as xplicit variable and so as to obtain a relationship for dy/dx or $dy/d\eta$, where η is linear multiple of x and $y = (b/c)\int^t x(\tau)\,d\tau$ or equivalently $x = (c/b)(dy/dt)$.

In obtaining Eq. (6.11–1) from a specific circuit, it is often convenient from the point of manipulation to write the small-signal transfer function from the rolled-source output, say $X_2(p)$, to the controlling variable, say $X_1(p)$, and then, multiplying out, to convert this to its equivalent large-signal nonlinear differ- l equation. A subsequent illustrative example will make this point clear. This sfer function should have the form

$$X_1(p) = X_2(p)\frac{pc}{p^2 + ap + b} \tag{6.11–2}$$

$$p^2X_1(p) + apX_1(p) + bX_1(p) = cpX_2(p),$$

re $X_1(p) = \mathscr{L}[x(t)]$ and $X_2(p) = \mathscr{L}[f(x)]$. Taking the inverse Laplace transform mmediately obtain Eq. (6.11–1).

Once Eq. (6.11–1) is obtained, we integrate once and divide through by c, the iplier of $f(x)$, to obtain

$$\frac{1}{c}\frac{dx}{dt} + \frac{a}{c}x + \frac{b}{c}\int^t x\,d\tau = f(x). \tag{6.11–3}$$

We now remove the integral term by defining a new variable,

$$y = \frac{b}{c}\int x\,d\tau \qquad \text{or} \qquad \frac{dy}{dt} = \frac{bx}{c}. \tag{6.11–4}$$

The dx/dt term is now removed by noting that

$$\frac{dx}{dt} = \frac{dx}{dy}\frac{dy}{dt} = \frac{dx}{dy}\frac{bx}{c}; \tag{6.11–5}$$

efore,

$$\left(\frac{b}{c^2}x\right)\frac{dx}{dy} + \frac{a}{c}x + y = f(x). \tag{6.11–6}$$

Now unless $b/c^2 = 1$, we define a new variable $\eta^2 = (b/c^2)x^2$ so that the slope has a unity multiplier and so that η and y have the same dimensions. We now the modified version of Eq. (6.11–6) as a solution for $dy/d\eta$,

$$\frac{dy}{d\eta} = \frac{\eta}{y - [f(x) - (a/c)x]}. \tag{6.11–7}$$

With this method it is convenient always to factor out a minus sign and to place shown in Eq. (6.11–7).

Now Eq. (6.11–7) presents the incremental change in y for an incremental change in η in terms of η, y, x, and $f(x)$. Since η is a linear function of x and y is an integral of x, we have as variables the controlling variable (or a linear multiple of it) and the integral of this variable (or a multiple of it). The two final variables are always connected via an integration or a differentiation and always have identical units.

At this point we lay out a set of axes with y as the ordinate and η as the abscissa and with *equal* increments in each direction. The coordinate system is known as the "phase plane." Now η and y in Eq. (6.11–7) are merely distances in this plane. If we assume values for x (and hence η), we may find a corresponding value for $f(x) - (a/c)x$ treated as a function of η. This function may be plotted in the y vs. η plane. Figure 6.11–1 illustrates Eq. (6.11–7).

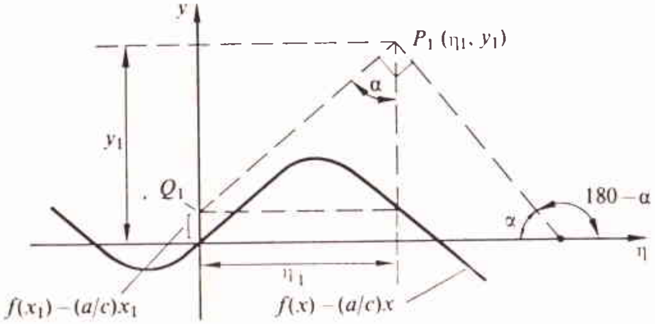


Fig. 6.11–1 Phase plane construction.

From any arbitrary point (η_1, y_1) in the phase plane, say P_1 in Fig. 6.11–1, we drop a perpendicular until it crosses the $f(x) - (a/c)x$ curve. At this point we travel horizontally until we reach the y-axis. From Fig. 6.11–1 it is clear that

$$\tan\alpha = \frac{\eta_1}{y_1 - [f(x_1) - (a/c)x_1]}$$

$$= -\frac{dy_1}{d\eta_1}.$$

However, $\tan(180° - \alpha) = -\tan\alpha$; hence

$$\frac{dy_1}{d\eta_1} = \tan\,(180° - \alpha).$$

This means that the gradient of y with respect to η at any point in the phase plane is given by the perpendicular to the radius from the equivalent point Q_1 to the point P_1 in question. Thus from any arbitrary starting point we may trace out a path by connecting successive arcs of circles swung about successive points on the y-axis.

Several points are worth noting before an illustrative example is undertaken.

Since we are working around an assumed Q-point, the $f(x) - (a/c)x$ curve can always be normalized so that it passes through the origin. Furthermore, unless it s through the origin with an upward slope to the right, that is, unless $f(x) >$ for small positive values of x, there is no sense in proceeding, because oscilla- are not possible [$f(x) > (a/c)x$ for small positive values of x implies that the gain exceeds unity and that the right-half plane poles exist in the system].

ecause the construction is both Q-point and circuit sensitive, any change in the it or the Q-point requires a new construction.

As an illustrative example consider the circuit of Fig. 6.10–1(b) with $R_2 = R_4$ $C_2 = C_4$ and the $f(i)$ vs. i characteristic as shown in Fig. 6.11–2.

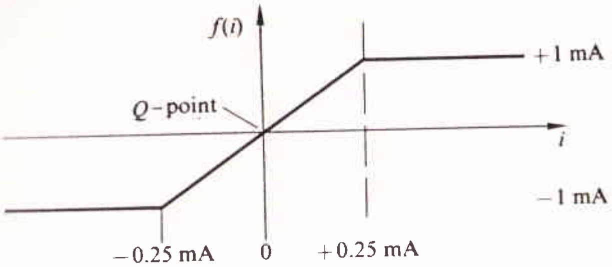


Fig. 6.11–2 Controlled-current-source characteristic.

From the circuit,

$$\frac{X_1(p)}{X_2(p)} = \frac{I_1(p)}{I_2(p)} = \frac{Y_4(p)}{Y_2(p) + Y_4(p)}. \tag{6.11–8}$$

For the specific case where $R_2 = R_4$, $C_2 = C_4$, and $\omega_0 = 1/R_2C_2$, the transfer ction reduces to

$$\frac{I_1(p)}{I_2(p)} = \frac{p\omega_0}{p^2 + 3\omega_0 p + \omega_0^2}; \tag{6.11–9}$$

ce $a = 3\omega_0$, $b = \omega_0^2$, and $c = \omega_0$.

Since Eq. (6.11–9) is of the form of Eq. (6.11–2) and since $b/c^2 = \omega_0^2/\omega_0^2 = 1$, we y write dy/di directly in the form of Eq. (6.11–7):

$$\frac{dy}{di} = -\frac{i}{y - [f(i) - 3i]}. \tag{6.11–10}$$

now assume variations of i *about the Q-point* and calculate $[f(i) - 3i]$ vs. i as tted in Fig. 6.11–3.

It is intuitively obvious that this circuit should oscillate: from the admittance nsfer function at $p = j\omega_0$ it follows that the network loss is $\frac{1}{3}$, while the small- nal current gain around the Q-point is 4; hence the loop gain exceeds unity.

Figure 6.11–3 also indicates the closed trajectory of the steady-state operating th. This path was sketched out with a compass in about three minutes from the itrary starting point of $y = 0$, $i = 0.50$ mA.

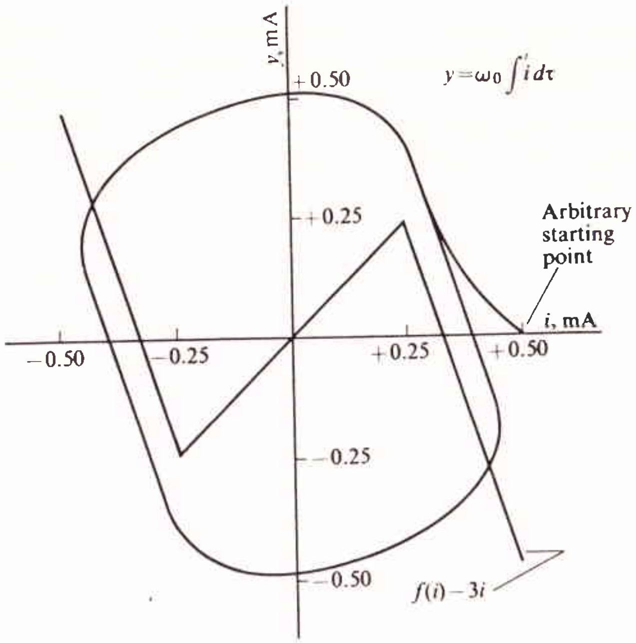


Fig. 6.11–3 Phase plane trajectory for circuit of Fig. 6.10–1(b) with $f(i)$ of Fig. 6.11–2.

Once we have this phase plot, the question arises as to what we can say about the waveshape of i or y vs. t. Since y and i are connected via an integration or differentiation, if one of them is sinusoidal then the other must be cosinusoidal. Therefore, since the scales are identical in both directions, it follows from what we know of Lissajous figures that if i were sinusoidal then the path would be circular. Therefore, the extent to which the actual trajectory departs from a circle gives us an indication of the nonsinusoidalness of the waveshape.

Before we look for further waveshape information we should consider the possible differences among the various available signals. Since $y = \omega_0 \int i\,d\tau$, then y is proportional to the voltage across C_4. Certainly the effect of the integration (a low-pass operation) should reduce the harmonic content in y with respect to that in i. Therefore, grounding the capacitor C_4 and taking the output across it should lead to a "better" waveshape than grounding R_4 and taking the output across it.

Whatever waveshape we desire, if we had the slope with respect to time for every point (as well as the corresponding value at that point), then we would be able to make a reasonable sketch of the waveshape with respect to time.

From our original equations we may write or derive both dy/dt and di/dt:

$$\frac{dy}{dt} = \omega_0 i, \qquad \frac{di}{dt} = -\omega_0\{y - [f(i) - 3i]\}.$$

These equations mean that the slope of y with respect to time is always proportional to the corresponding value of i or to the horizontal distance from the vertical s to the point on the trajectory. At the same time, the slope of i with respect to ie is proportional to the negative of the vertical distance between the trajectory l the $f(i) - 3i$ characteristic. When the trajectory crosses the $f(i) - 3i$ characistic, di/dt goes through zero; when the trajectory crosses the y-axis, dy/dt goes zero.

If one is attempting to plot $y(t)$, then the easiest approach would seem to be to iw a circle with center at the origin and radius equal to $y_{\max}$. Then sketch out efully on a separate sheet of graph paper about $1\frac{1}{2}$ cycles of a cosine wave of itrary period. So long as the trajectory lies on the circle, then $y(t)$ is following the e wave. As the trajectory falls away from the circle, the slope of $y(t)$ decreases m the slope of the cosine wave. If the slope of the cosine as it passes through zero aken as the standard, then the slope of $y(t)$ at any point is merely $i_{\text{point}}/y_{\max}$ times standard slope.

The slope-tracking approach will break down near the negative peak in y. owever, even if the negative maximum is different in amplitude from the positive ximum, the trajectory in the vicinity of the peak is still nearly an arc of a circle; ice the negative tip of $y(t)$ will be sinusoidal in shape. With this information it is a aightforward task to sketch out $y(t)$. Since $v_{C4}(t) = R_4 y(t)$, this sketch provides the

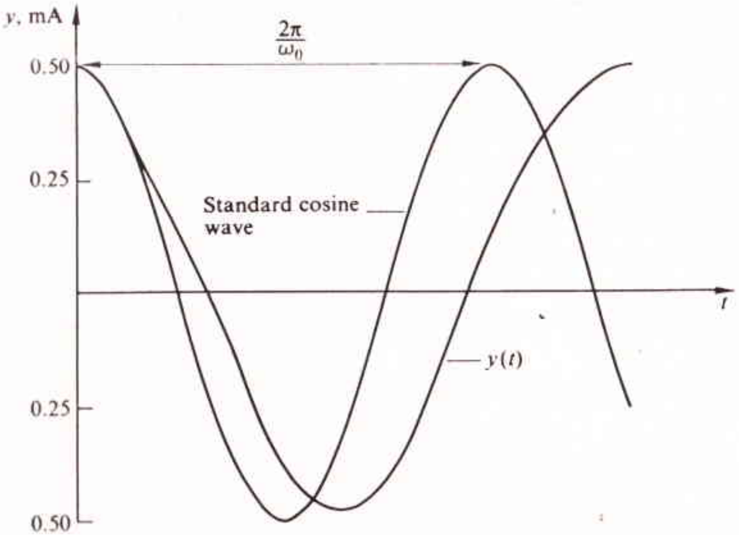


Fig. 6.11-4 Sketch of $y(t) = v_{C4}(t)/R_4$ for the case shown in Fig. 6.11–3.

tput voltage within a constant. Figure 6.11–4 compares $y(t)$ with a standard sine ve of the same peak amplitude and of period ω_0. Note that, in spite of the reasonly strong overdrive, the waveshape across C_4 is a good approximation of a sine ve, but a sine wave of a frequency at approximately 80% of ω_0.

As the overdrive is reduced, the slope of $f(i) - 3i$ through the origin is reduced til in the limit it becomes horizontal. At this point i would be sinusoidal with a

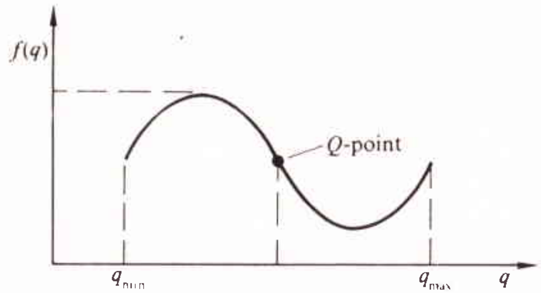
---

peak amplitude equal to the breakpoint value of i (assuming a symmetrical characteristic is maintained as the slope is reduced). Any further reduction in the slope will cause the oscillations to cease.

As the overdrive is reduced, not only do the waveshapes become more sinusoidal but their frequencies increase toward ω_0. In the next section we shall see how the frequency of an oscillator is affected by distortion in its waveshape and hence why the overdrive (with its large distortion terms) should cause a frequency depression.

6.12 THE DISTORTION–OPERATING FREQUENCY RELATIONSHIP†

Almost any practical sinusoidal or nearly sinusoidal oscillator may be considered in a form similar to one of the circuits shown in Fig. 6.10–1. That is, it may be considered as a combination of a network and a controlled source. In general, we have been neglecting active circuit reactances; however, even if they occur it is usually possible to segregate them and then to treat them as part of the network. Thus if we wish we can consider all of our oscillations as a network plus a nonreactive controlled source.



$$\oint f(q)\,dq = 0$$

$$\oint q\,d[f(q)] = 0$$

Fig. 6.12-1 General nonreactive controlled-source characteristic.

Whatever the combination of variables,‡ a nonreactive active element driven by a periodic waveshape (we assume steady-state operation, but make no initial assumptions about the waveshape other than that it is periodic) must have zero net area when integrated around a complete cycle. Figure 6.12–1 illustrates this property for the general controlling variable q and controlled variable $f(q)$.

If the active device were reactive, then for any given amplitude swing for q the result of the hysteresis in the characteristic would be a constant area and our approach would still be valid although slightly more complicated. Therefore, as a simplification we separate the device reactances initially and assume no device hysteresis.

---

† The relationship between distortion and operating frequency of an oscillator was first presented by Groszkowski in the 1920's. His more recent book, *Frequency of Self Oscillations* (Macmillan, New York, 1964), summarizes his earlier work and presents results for more complicated cases than are presented here.

‡ Controlled-source possibilities include current-controlled voltage or current sources and voltage-controlled voltage or current sources; network functions include impedances, admittances, and voltage and current ratios.

Now instead of taking the integration from $q_{\min}$ to $q_{\max}$ and back, we could take it over a complete period of the input cycle, in which case we would find

$$\oint f(q)\,dq = \int_0^T \left[f(q)\frac{dq}{dt} \right] dt = 0. \tag{6.12–1}$$

Furthermore, since q and hence $f(q)$ are periodic variations about a Q-point, each of them can be expanded in a Fourier series starting from a fundamental component and going as high as is necessary:

$$q = \sum_{n=1}^{\infty} Q_n \sin(n\omega t + \psi_n), \tag{6.12–2}$$

$$\frac{dq}{dt} = \sum_{n=1}^{\infty} n\omega Q_n \cos(n\omega t + \psi_n), \tag{6.12–3}$$

$$f(q) = \sum_{m=1}^{\infty} F_m \sin(m\omega t + \phi_m). \tag{6.12–4}$$

If Eqs. (6.12–3) and (6.12–4) are substituted in Eq. (6.12–1) and the integration is performed, then all terms except those containing $\sin(\psi_n - \phi_m)$, where $m = n$, vanish and we obtain

$$\sum_{n=1}^{\infty} \frac{Q_n F_n n\omega T}{2} \sin(\psi_n - \phi_n) = 0. \tag{6.12–5}$$

Equation (6.12–5) indicates an active-device-imposed relationship between the harmonics of the controlling and controlled quantities and the phase angle between them.

The passive circuit also imposes a different relationship between these quantities. These two relationships must coexist. In general, this means that the operating frequency of the oscillator must shift until the two relationships can be mutually satisfied.

If the transfer function between $f(q)$ and q through the passive network of the oscillator is $H(p)$, then it is readily shown that

$$F_n\, \mathcal{I}m\, \mathrm{H}(jn\omega) = Q_n \sin(\psi_n - \phi_n),$$

where $\mathcal{I}m\, H(jn\omega)$ is the imaginary part of $H(jn\omega)$. With this result, Eq. (6.12–5) simplifies to

$$\sum_{n=1}^{\infty} nF_n^2\, \mathcal{I}m\, H(jn\omega) = 0 \tag{6.12–6}$$

or

$$\mathcal{I}m\, H(j\omega) = -\sum_{n=2}^{\infty} n\, \mathcal{I}m\, H(jn\omega)\frac{F_n^2}{F_1^2}. \tag{6.12–7}$$

Here we see that the operating frequency must shift from ω_0 [the point at which $H(j\omega)$ is real] to a point where $H(j\omega)$ is sufficiently reactive to satisfy Eq. (6.12–7). It is also apparent that this shift in frequency is minimized if the harmonic components in $f(q)$ are small.

Up to this point we have not made any approximations whatsoever. In many cases, it is convenient to make two types of simplifying approximations. In one, we assume a simple driving function for $q(t)$ in order to ease the calculation of the harmonics of $f(q)$. In the other, we make some simplifying assumptions about the network pole-zero patterns in order to facilitate the calculation of the $\mathcal{I}m\, H(jn\omega)$.

As a numerical example, consider the circuit of Fig. 6.10–1(a) as expanded in conjunction with Figs. 6.10–4, 6.10–5, and 6.10–6. For the given drive, let us find the first-order effect of the circuit Q_T on the operating frequency. To accomplish this we assume that Q_T is sufficiently high so that i remains essentially sinusoidal as Q_T varies.

For this case the variable Q_n becomes I_n, the variable F_n becomes V_n, and the variable $\mathcal{I}m\, H(jn\omega)$ becomes $\mathcal{I}m\, Y(jn\omega)$, where

$$Y(p) = \frac{p/L}{p^2 + \dfrac{r}{L}p + \dfrac{1}{LC}}, \qquad r = r_1 + r_2,$$

and

$$Y(jn\omega) \approx Y(jn\omega_0) \approx \frac{-jn}{\omega_0 L(n^2 - 1)} \qquad \text{for } n \geq 2 \text{ and high } Q_T;$$

thus

$$\mathcal{I}m\, Y(jn\omega) \approx \frac{-n}{\omega_0 L(n^2 - 1)}.$$

In addition,

$$\mathcal{I}m\, Y(j\omega) = \frac{(\omega/L)(\omega_0^2 - \omega^2)}{(\omega_0^2 - \omega^2)^2 + (\omega r/L)^2} \approx -\frac{2\Delta\omega Q_T^2}{\omega_0^2 L} \qquad \text{for high } Q_T,$$

where $\Delta\omega = \omega - \omega_0$ is the departure in operating frequency from ω_0. Thus for this example we obtain, from Eq. (6.12–7),

$$\frac{\Delta\omega}{\omega_0} \approx -\frac{1}{2Q_T^2} \sum_{n=2}^{\infty} \frac{n}{n^2 - 1} \left(\frac{V_n}{V_1} \right)^2. \tag{6.12–8}$$

Now $(V_n/V_1)^2$ may be evaluated for any particular steady-state operating point by using Fig. 4.2–3. Table 6.12–1 lists the approximate harmonic ratios for the first five harmonics on the assumption of operation around the operating point found in Section 6.10.

Table 6.12–1

$V_2/V_1 \approx 0.0$
$V_3/V_1 \approx 0.32$
$V_4/V_1 \approx 0.0$
$V_5/V_1 \approx 0.15$
$I_1 \approx 3.8\text{ mA},\ V_1 \approx 7.6\text{ V}$

With these values of V_n/V_1 and with $Q_T = 12$, we find that

$$\begin{aligned}\frac{\Delta\omega}{\omega_0} &\approx \tfrac{1}{288}[(\tfrac{3}{8} \times 0.32^2 + \tfrac{5}{24} \times 0.15^2)] \\ &\approx 1.5 \times 10^{-4},\end{aligned}$$

or that the effect of the harmonics has been to decrease the operating frequency by 0.015%. With a smaller Q_T there would be more harmonics in i and thus more uncertainty in estimating V_n. A smaller Q_T would certainly also lead to a greatly increased $\Delta\omega$.

Equations (6.12–7) and (6.12–8) indicate that the operating frequency of an oscillator may be shifted by injecting harmonics of the oscillator frequency or by varying the phase of these injected harmonics. This leads directly to the possibility of synchronization with an external signal either at its fundamental or any of its harmonics.

From this point of view we can see why the oscillator explored in Section 6.11 operated so far below its nominal frequency of operation. To begin with, the oscillator was badly overdriven, so that the harmonic content was high; but to make matters worse, the passive circuit transfer function

$$H(p) = \frac{p\omega_0}{p^2 + 3p\omega_0 + \omega_0^2}$$

neither attenuated the harmonics rapidly with respect to the fundamental nor provided a rapid phase shift with frequency in the neighborhood of ω_0. Thus a large decrease in frequency was necessary to bring Eq. (6.12–7) back into balance.

PROBLEMS

6.1 a) In the circuit of Fig. 6.P–1, $A_1 = A_2 = A_3$ and all sources are linear. Find the values of C and $A_{\min}$ required if this circuit is to be on the verge of sinusoidal oscillation at 20 kHz. How would the circuit be affected if the signs of one, of two, or of three of the generators were reversed?

b) With the same value for C, assume that $A_2 = A_3 = 4$ while amplifier 1 has the following nonlinear relationship between its input rms sinusoidal voltage and its output sinusoidal voltage:

$$|V_o(\omega)| = 1 - \frac{|V_i(\omega)|^2}{2}.$$

Find the sinusoidal output voltage $v_3(t)$. (The nonlinear relationship is assumed to have a time constant that is long in comparison to a cycle of the sine wave.)

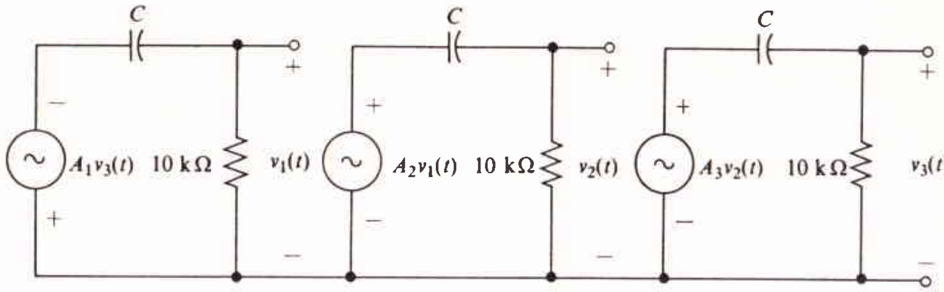


Figure 6.P–1

6.2 If all of the 10 kΩ resistors in Problem 6.1(b) increase 10% with age, what would be the new values of frequency and output amplitude? If R did not vary, but A_2 decreased from 4 to 3, what would be the effect on frequency and amplitude?

6.3 A possible sine-wave oscillator circuit is shown in Fig. 6.P–2. Will it work? If so, find the frequency and the amplitude of v_x. If not, explain why not. Assume that the AVC block does not load Q_2 and that $\alpha \approx 1$ for the transistors.

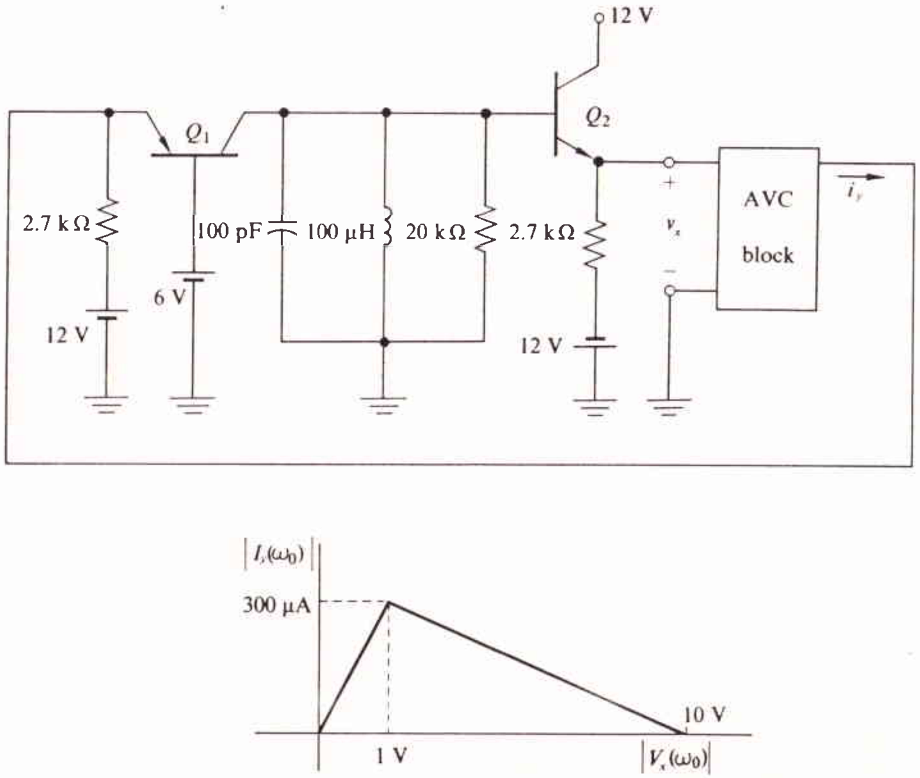


Figure 6.P–2

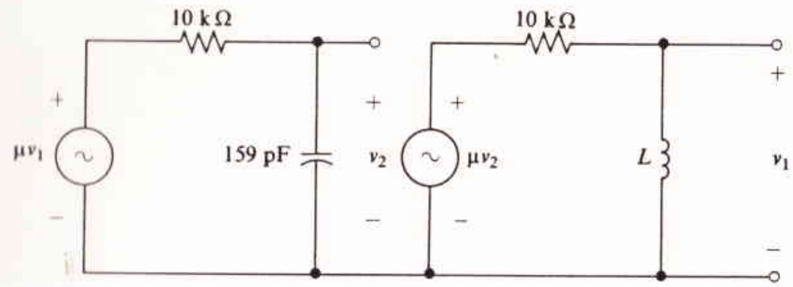
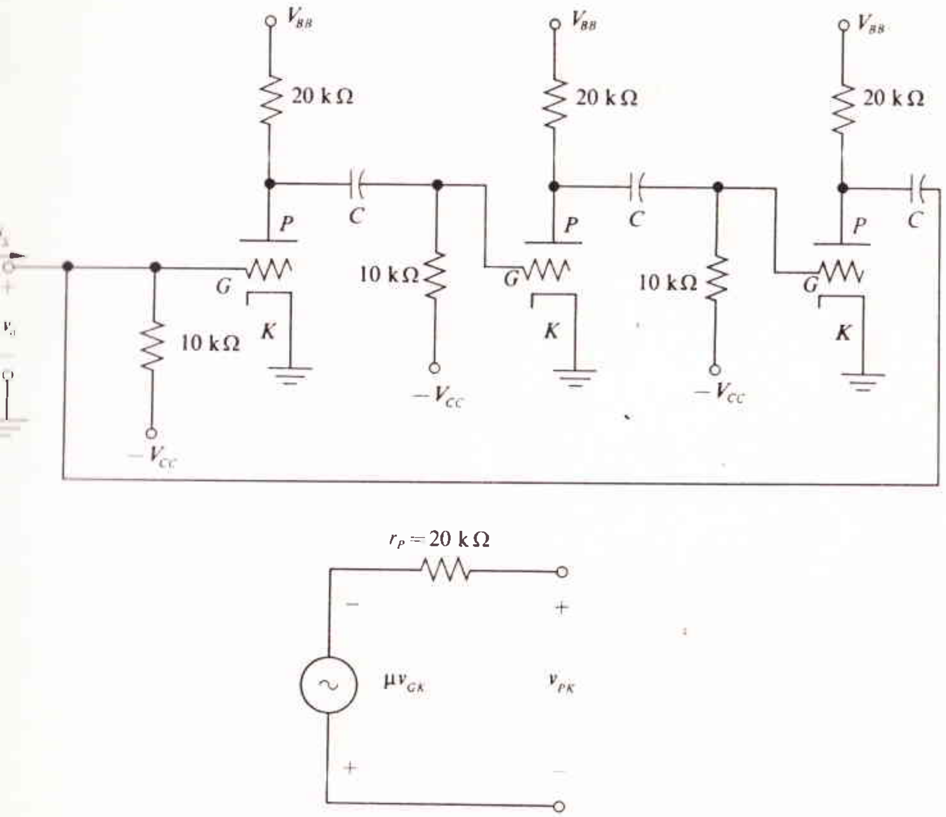


Figure 6.P–3

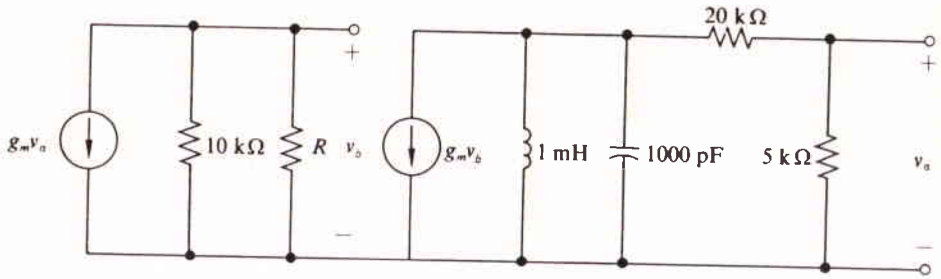
4 For the circuit of Fig. 6.P–3, find μ_{min} for oscillations. Find L for 100 kHz oscillations.

5 For the circuit of Fig. 6.P–4, indicate the loci of the poles of $V_a(p)/I_S(p)$ as μ varies, assuming identical μ's and that a current generator, i_S, is applied at the terminals at v_a. What is the minimum value of μ for sinusoidal oscillations? What value of C will cause sinusoidal oscillations at 10 kHz?



Small-signal triode model

Figure 6.P–4



$g_m = 1000\mu℧$ with $v_b = V_b \sin \omega t$

$R = \dfrac{10\text{ k}\Omega}{V_b^2}$

Figure 6.P–5

6.6 In the circuit of Fig. 6.P–5, R is a temperature-sensitive resistor with a long time constant. Find v_b. What is the Q of the tuned circuit? What are the form and magnitude of v_a?

6.7 In the circuit of Fig. 6.P–6(a), assume that the box B has $R_{in} = 25$ kΩ, no phase shift, and the amplitude characteristic shown in Fig. 6.P–6(b). Letting $N = 5$, find the frequency and amplitude of v_1. (Assume identical tubes and the model shown in Fig. 6.P–4.)

6.8 With a pole-zero pattern of the type shown in Fig. 6.1–7(c), produced by cascading three identical isolated RC circuits, show that $\omega_0 = \sqrt{3}/RC$ and $A_{min} = -8$. Indicate a possible circuit configuration for this case.

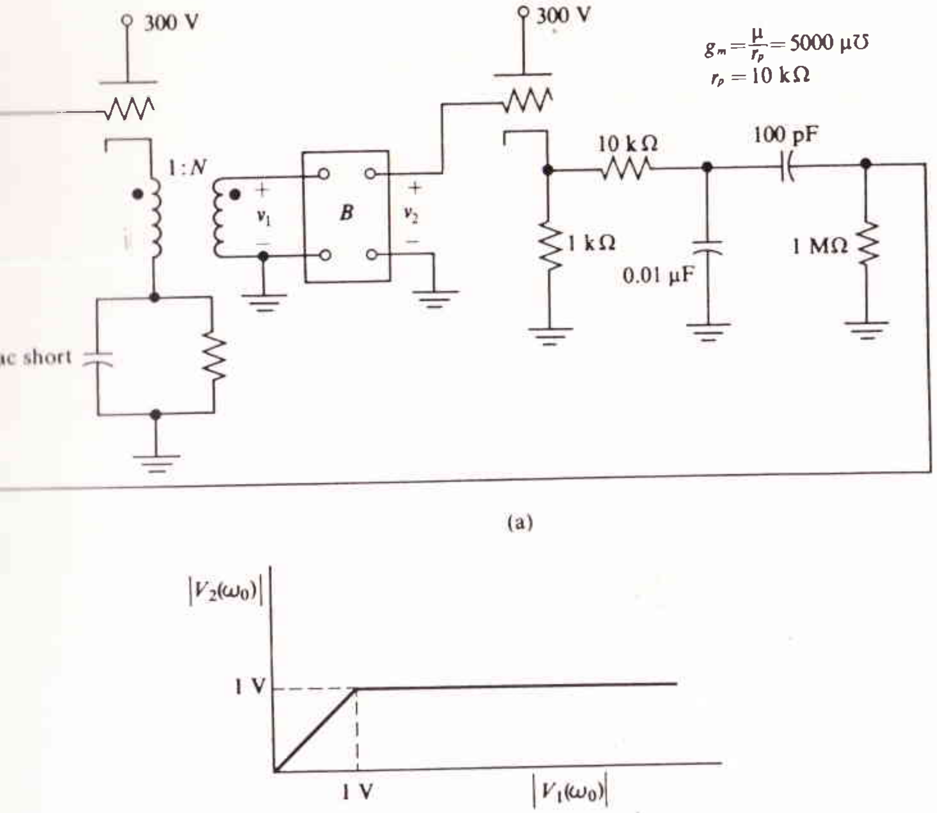
6.9 Assume that three identical, isolated RC circuits are used to produce a network for a feedback oscillator with a root locus pattern as shown in Fig. 6.1–7(d). Prove that for sinusoidal oscillations $\omega_0 = 1/\sqrt{3}RC$ and that $A_{min} = -8$. Indicate a possible transistor circuit configuration for this case. What values would ω_0 and A have if the isolated center section were readjusted to have twice the time constant of the two outside sections?

6.10 Derive Eq. (6.9–8). Assuming that $R_1 = R_2 = R_4 = R$, $C_1 = C_2 = C$, and $R_3 = 2.010R$, relate the transfer function of the Wien bridge to the pole-zero pattern of Fig. 6.1–7(a) and determine the sign of A and the magnitude of A_{min} necessary if the circuit is to oscillate. Repeat, letting $R_3 = 2.0010R$. What is the S_F for each case?

6.11 Derive Eq. (6.9–3) for the Meacham bridge oscillator shown in Fig. 6.9–1(a). Let $\omega_0 = 10^6$ rad/sec, $R_3 = 100\,\Omega$, $R_2 = 50\,\Omega$, $R_1 = 200\,\Omega$, and $Q = 10^4$ (a quartz crystal is used as a series tuned circuit). Find r for oscillation if the available A is 200. What is the sign of A? With these values, assume that the phase shift in the amplifier shifts by 0.10 rad. Find the shift in frequency from ω_0.

6.12 For the circuit shown in Fig. 6.P–7, determine the output collector current components at 1050 kHz and at 2100 kHz. Find the voltage across L_2 at 1050 kHz. The transistor is silicon and $\alpha = 0.99$; $C_2/C_1 = 40$, $L_2 = 150\mu$H, $Q_{L2} = 50$, and $M_{12} = 4\,\mu$H. The tuned circuit is tuned to 1050 kHz.

6.13 Find the frequency and amplitude for the voltage $v_o(t)$ for the circuit shown in Fig. 6.P–8. Find the total harmonic distortion in the output. The transistor is silicon with $\alpha = 0.98$ and $I_{ES} = 3 \times e^{-30}$ mA; $L_1 = 10\,\mu$H, $L_2 = 250\,\mu$H, $M_{12} = 25\,\mu$H, $C_1 = 200$ pF, $C_2 = 0.04\,\mu$F, and $R_2 = 100$ kΩ.



(a)

(b)

Figure 6.P–6

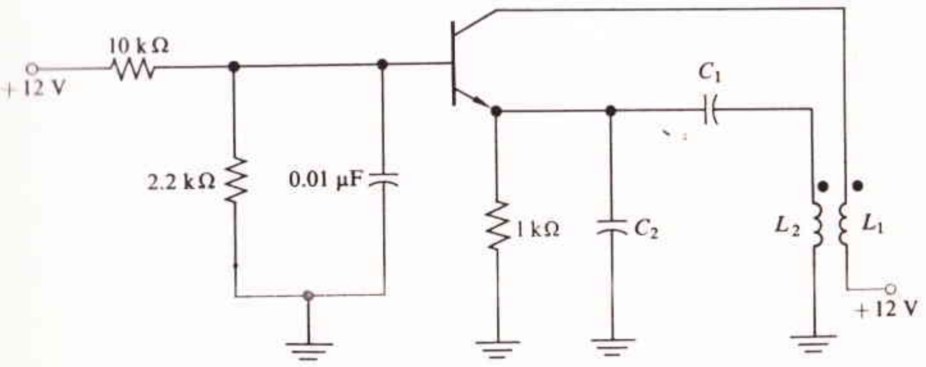


Figure 6.P–7

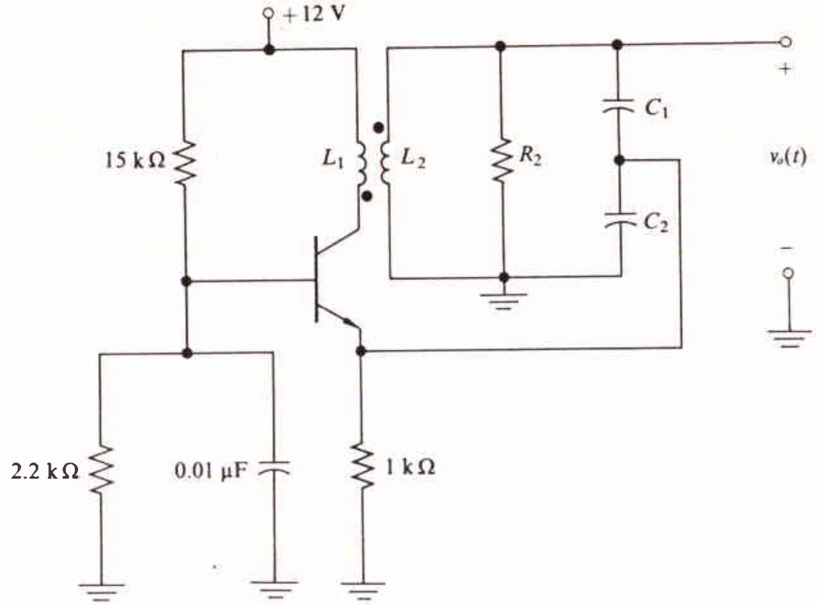


Figure 6.P–8

6.14 Is squegging possible in the circuit of Fig. 6.P–7? If so, at what sinusoidal frequency will it occur? If not, then would it be possible if the tuned circuit's Q were reduced to 20?

6.15 A simplified version of a commercial RC oscillator is shown in Fig. 6.P–9. The network transfer function is most easily found by superposition, that is, by (1) calculating the voltage at the input to amplifier 1 as a function of $v_o(t)$ assuming that amplifier 2 has zero output impedance and zero output voltage, and then (2) calculating the voltage at the input to amplifier 1 assuming that amplifier 1 has zero output impedance and voltage while the voltage output of amplifier 2 is $Nv_o(t)$. These two results are then combined to produce the input to amplifier 1 (assumed as having an infinite input impedance) as a function of $v_o(t)$. If the two operational amplifiers have equal voltage gains of A, then show that stable sinusoidal oscillations occur with $\omega_0 = 1/RC$ and $N = 2 + (3/A)$.

If $A = 100$, $r = 300\ \Omega$, and $r_v = 450 \exp[-65P]$, where P is the ac power dissipated in r_v, what is the value of $v_o(t)$ at equilibrium? What is the value of r_v for this case? If A were increased to 400 for each amplifier, how much would $v_o(t)$ increase? How does the indirect frequency stability S_F of this circuit compare with that of a Wien bridge oscillator?

6.16 All transistors in the circuit of Fig. 6.P–10 are silicon with $\alpha \geq 0.99$, Q_4 and Q_5 are identical, $Q_{L1} = Q_{L3} = 100$, $\omega_0 L_1 = 1000\ \Omega$, $M/L_1 = 1/20$, $C_1/[C_1 + C_2] = 1/40$, and L_3C_3 is tuned to the same frequency as L_1, C_1, and C_2. The total load on the L_3C_3 tuned circuit including its own losses is 20 kΩ. Plot $G_m(x)$ vs. x for this circuit, assuming an ideal envelope detector. Find the stable oscillation amplitude existing across L_1. Find I_k for this amplitude. Find the changes in this amplitude caused by a ±10% change in the +6 V supply. Find S_F for this oscillator.

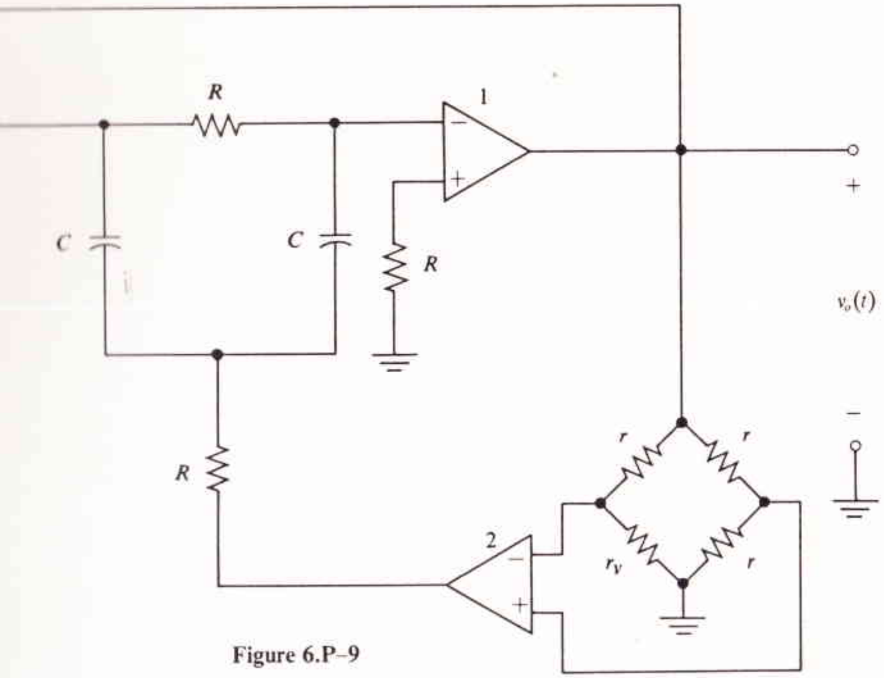


Figure 6.P–9

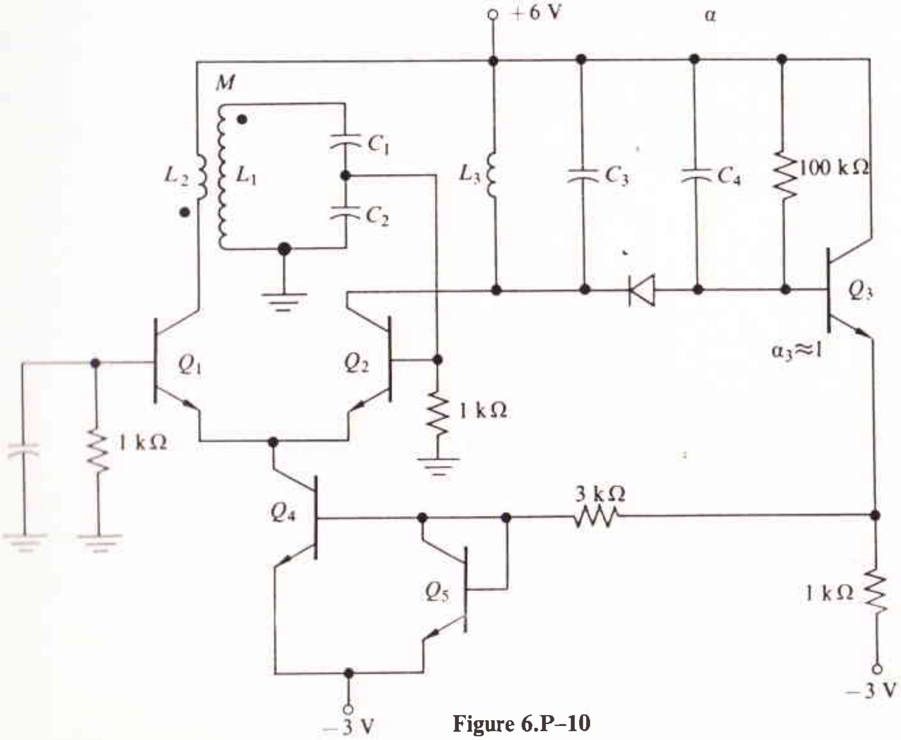
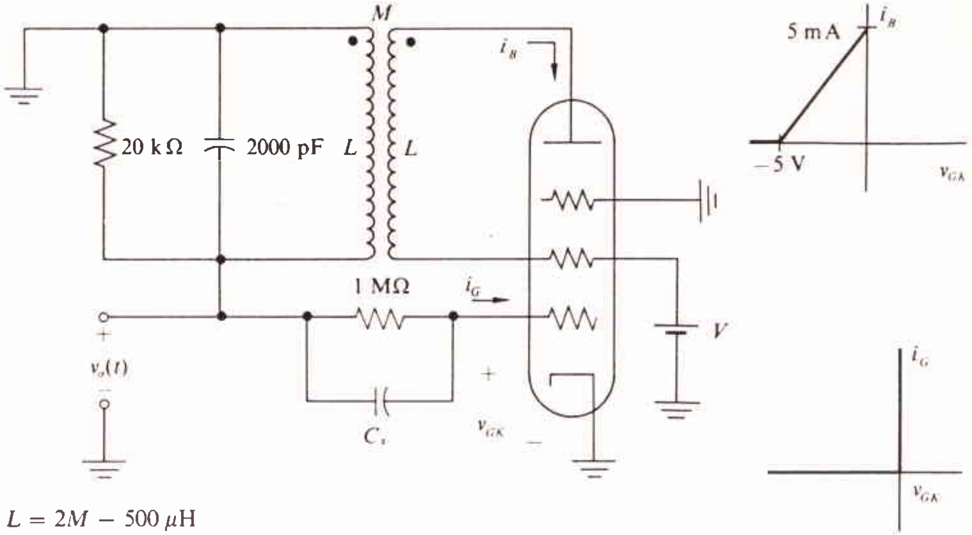


Figure 6.P–10



$L = 2M - 500\ \mu\text{H}$

Figure 6.P–11

6.17 For the circuit of Fig. 6.P–11, find frequency and amplitude of v_o assuming that the loss in the 1 MΩ resistor is negligible, that C_x is an *RF* bypass, and that the grid diode is ideal as shown.

6.18 For the circuit of Fig. 6.P–12, find the voltage magnitude and frequency across R_L. (Neglect losses in coil.) Repeat, reducing V_{CC} to 12 V. Is squegging possible in the circuit shown in Fig. 6.P–12? At what frequency does it occur? The transistor is silicon.

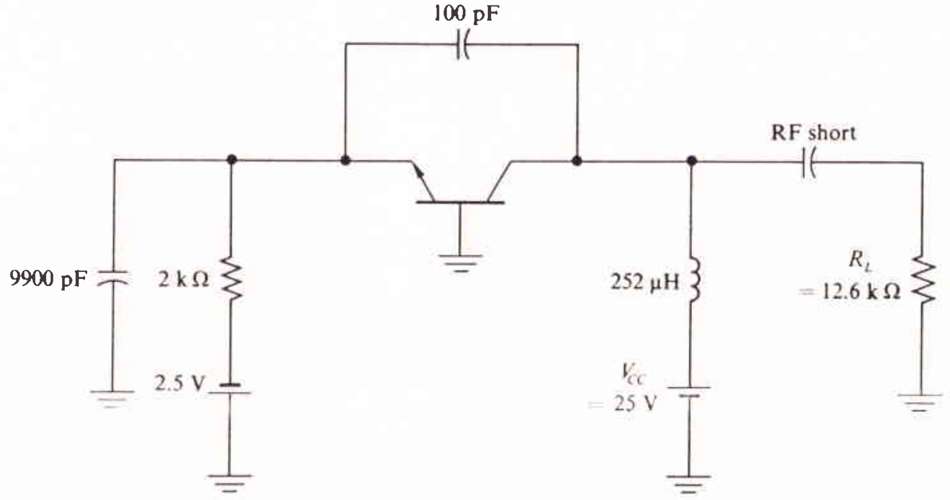


Figure 6.P–12

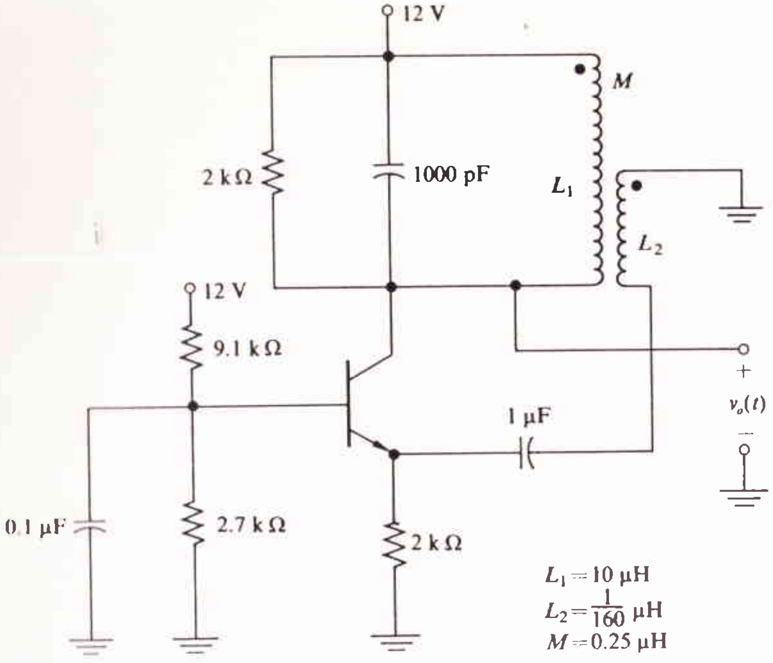


Figure 6.P–13

For the circuit shown in Fig. 6.P–13, determine an expression for $v_o(t)$.

A sine-wave oscillator constructed from a commercial differential pair–current source combination is shown in Fig. 6.P–14.

a) Find I_k.
b) Choose R_2 so as to make $v_{B2}(t) = 0.104 \cos \omega_0 t + V_{dc}$.
c) Find the voltage across each tuned circuit and $v_o(t)$.
d) Find the THD in $v_o(t)$.

For the circuit of Fig. 6.P–15, $R = 10\,\Omega$, $C = 200$ pF, $L = 100\,\mu$H, and $v_D = (11 - i_D + 2i_D^2 + 3i_D^3)$ V, where v_D is measured in volts and i_D is measured in mA. Find the frequency and the amplitude of the fundamental component of $v_C(t)$. What is the total harmonic distortion in $v_C(t)$? How much frequency depression is caused by this distortion?

Assume $v_D = V_1 \cos \omega_0 t$ for the circuit shown in Fig. 6.P–16.

a) Plot I_{D1} vs. V_1 for the device shown.
b) Determine a numerical expression for the steady-state value of $v_D(t)$.

Find the approximate amplitude of $v_D(t)$ in the circuit shown in Fig. 6.P–17.

Plot the phase plane characteristic for the device characteristic and circuit shown in Fig. 6.P–18. Start the plot at $v_o = 0.0$ V, $i_T = 0.50$ mA, and let $r = 16$ kΩ, $\omega_0 L = 32$ kΩ, and $Q = 2$.

Plot the output waveshape for $v_C(t)$ in Fig. 6.P–18. Estimate the frequency depression for the actual output in comparison to $\omega_0 = 1/\sqrt{LC}$.

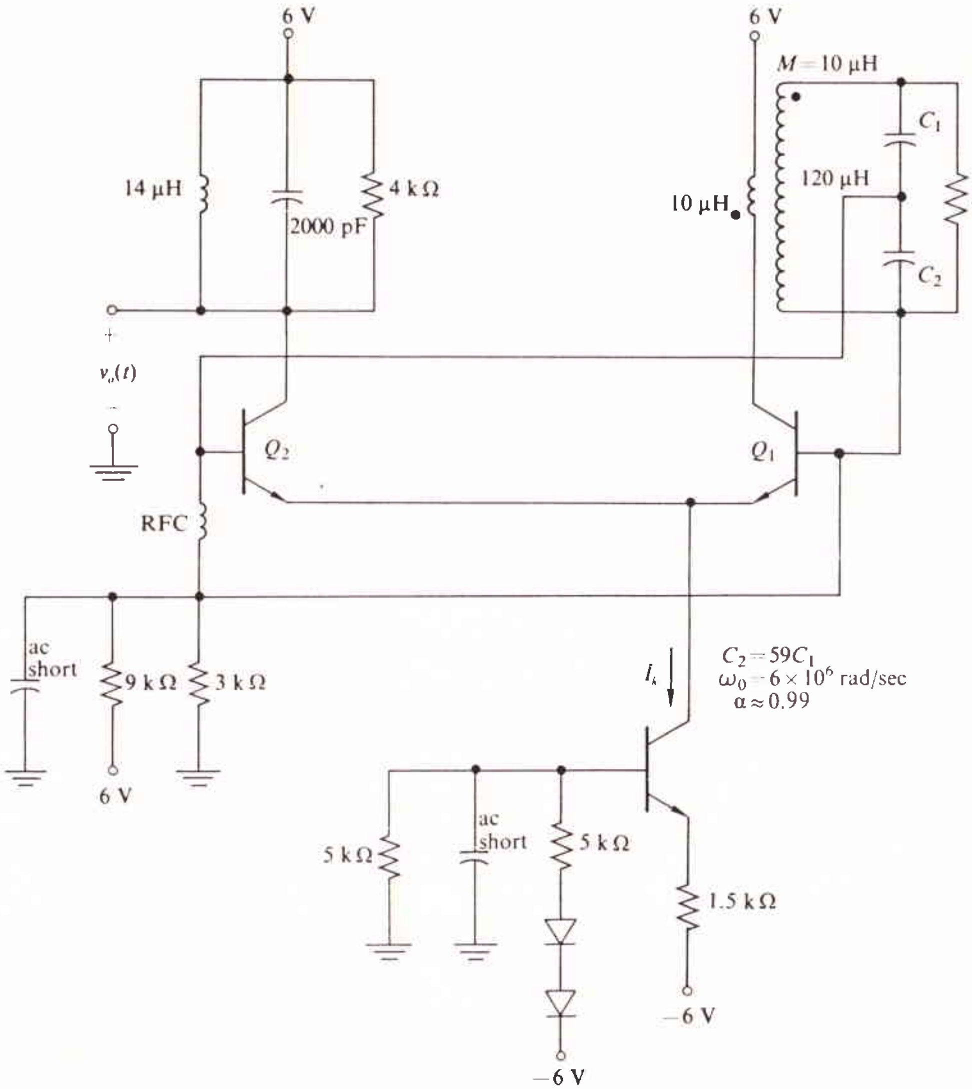


Figure 6.P–14

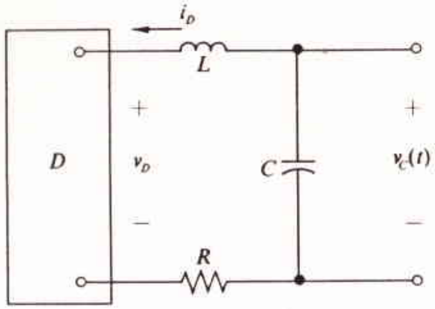


Figure 6.P–15

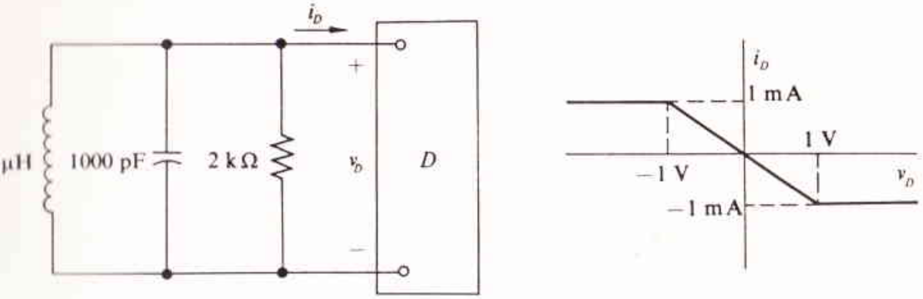
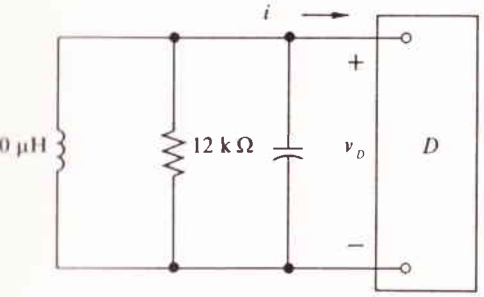


Figure 6.P–16



$\omega_0 = 6 \times 10^6$ rad/sec
$i_D = (-0.2 \times 10^{-3} v_D + 0.75 \times 10^{-6} v_D{}^3)A$

Figure 6.P–17

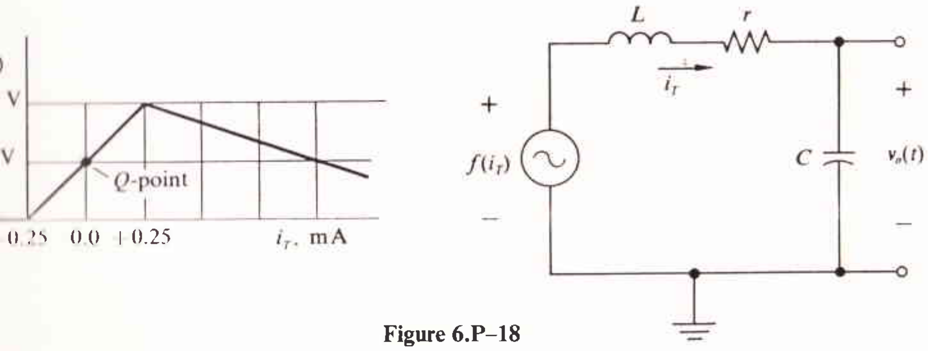


Figure 6.P–18

CHAPTER 7

MIXERS; RF AND IF AMPLIFIERS

7.1 THE SUPERHETERODYNE CONCEPT

In designing a receiver, one normally starts from the detector or demodulation circuitry and works in both directions. In this chapter we are interested in the circuits that lie in front of the demodulators.

As we shall see in subsequent chapters, most detector circuits do not work well in the presence of noise or interfering signals, and many of them do not work well, if at all, with input signal levels below several volts in amplitude. Since the desired signal may have a receiver input field strength in the microvolt/meter range while the total rms noise and interfering signal strengths available to the antenna may measure in the volt/meter range, it is apparent that one needs both gain and selectivity in front of the demodulator.

The real problems in designing a carrier frequency or RF amplifier for a fixed-carrier-frequency receiver are the following:

a) to control its front-end noise so as to keep it an adequate distance below the incoming signal level;
b) to control the active device nonlinearities so as to prevent signal distortion or unwanted signal interactions [cross modulation, for example]†;
c) to keep the resultant high-gain narrowband amplifier from becoming an oscillator. [If the amplifier gain is 120 dB (10^6 in voltage), then a feedback of $1/10^6$ of the output, with the proper phase relationship, to the input will cause a loop gain of unity and hence will produce an oscillator.]

Unfortunately for receiver designers, most receivers are not fixed-frequency units; hence one is forced to cope with the previously mentioned problems while simultaneously tuning this high-gain monster over some wide frequency range. In addition, one must be able to solve the difficult problem of designing the demodulator to have adequate and reasonably constant performance over the frequency band(s) in question. (Normal AM broadcasting has a 3/1 band, normal VHF television in the United States has a 4/1 band, and a "good" communications receiver may be expected to cover a range of more than 100/1 in frequency.)

---

† As we shall see shortly, cross modulation is the transfer of the modulating signal, or a distorted version of this signal, from one carrier to a neighboring carrier. In order for such a phenomena to be possible, certain types of device nonlinearities must be present.

Very early in the development of the radio communications business some people ecided that this situation was ridiculous and that the way to simplify things was to ontinue to design both the detector and the bulk of the gain and selectivity on a ed-frequency basis and to shift or translate the modulation from all desired in- oming signals to this new, fixed, "intermediate frequency" or IF. As with most new eas, it took some time for this innovation to be accepted; however, it has since so ompletely dominated the field that "straight-through" receivers are rarely seen cept in museums and an occasional very special situation.

Figure 7.1–1 illustrates the block diagram of a "superheterodyne" receiver. This the common name for the system containing an IF amplifier and a fixed-frequency etector.

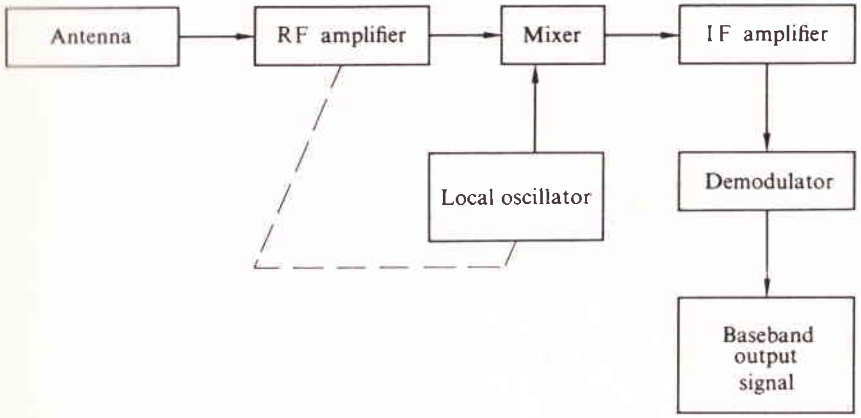


Fig. 7.1–1 Superheterodyne receiver.

The "mixer" utilizes the trigonometric identity that expands the product of o cosine terms into sum and difference frequencies:

$$[a(t)\cos At][b(t)\cos Bt] = \frac{a(t)b(t)}{2}[\cos(A - B)t + \cos(A + B)t]. \qquad (7.1\text{–}1)$$

us if $a(t)\cos At$ is the desired incoming signal while $b(t)\cos Bt$ is a constant- nplitude cosine signal supplied by the local oscillator, and if the selective IF nplifier is tuned to a radian frequency of $A - B$, then the IF output signal will be frequency-translated version of the incoming signal. So long as everything is done operly, such a frequency translation is modulation insensitive. That is, it is just effective for AM, FM, SSB, or any other type of modulation.

Like most innovations the superheterodyne concept does add some additional oblems to those already listed for the case of the fixed-frequency receiver. Some these additional problems are as follows:

the mixer and local oscillator must be designed and the local oscillator must be made to track all tuned circuits in front of the mixer;

e) since mixers almost always generate more noise than amplifiers and since by their very nature they must contain nonlinearities, one may find that criteria (a) and (b) still require a stage (or stages) of RF amplification in front of the mixer.

f) Some new types of interference are generated by the local oscillator–mixer combination that are not present in a straight-through operation.

The next section will deal with several general techniques for analyzing nonlinear circuits as mixers. Subsequent sections will deal with specific mixing and "converting" circuits (a self-oscillating mixer is called a converter) and with the design and analysis of RF and IF amplifier stages.

7.2 MIXER TECHNIQUES

As was pointed out in connection with Eq. (7.1–1), any multiplier followed by a proper bandpass filter will function as a mixer. On the other hand, since the local oscillator input has a constant amplitude, it is not necessary to have an ideal multiplier in order to make a useful mixer.

Chapter 8 will consider several general multiplier circuits. The emphasis in this section will be on techniques applicable directly to specific mixer circuits.

The two most common mixers in use today are the simple field effect transistor and the bipolar transistor. In both cases one applies the incoming signal and the local oscillator voltages so that they effectively add to the dc bias voltage to produce the total gate-source or base-emitter voltage. This signal is then passed through the device nonlinearity to create the desired sum and difference frequencies.

FET Mixers

Figure 7.2–1 illustrates several possible FET mixer circuits.

If a junction or MOS FET is biased so that the total excursions around the Q-point never cause it to leave the "constant current" (saturation region) or to turn on the gate-to-source junction, then the drain current is always approximately related to the gate-source voltage by a "square-law" characteristic, and the drain currents are given by

$$i_D = I_{DSS}\left[1 - \frac{v_1(t) + v_2(t) + V_{GS}}{V_p}\right]^2 \qquad (7.2\text{–}1)$$

and

$$i_D = \frac{\beta}{2}[v_1(t) + v_2(t) + V_{GS} - V_{Th}]^2, \qquad (7.2\text{–}2)$$

respectively. Where $v_{GS} = v_1 + v_2 + V_{GS}$, v_1 and v_2 are the local oscillator and RF input signals respectively, and V_{GS} is the dc gate-to-source bias voltage. These equations are illustrated in Fig. 7.2–2. If Eq. (7.2–2) for the MOS unit is chosen as an example and the voltage term is expanded, then

$$i_d = \beta\left\{v_1(t)v_2(t) + (V_{GS} - V_{Th})[v_1(t) + v_2(t)] + \frac{[v_1(t)]^2}{2} + \frac{[v_2(t)]^2}{2} + \frac{(V_{GS} - V_{Th})^2}{2}\right\}. \qquad (7.2\text{–}3)$$

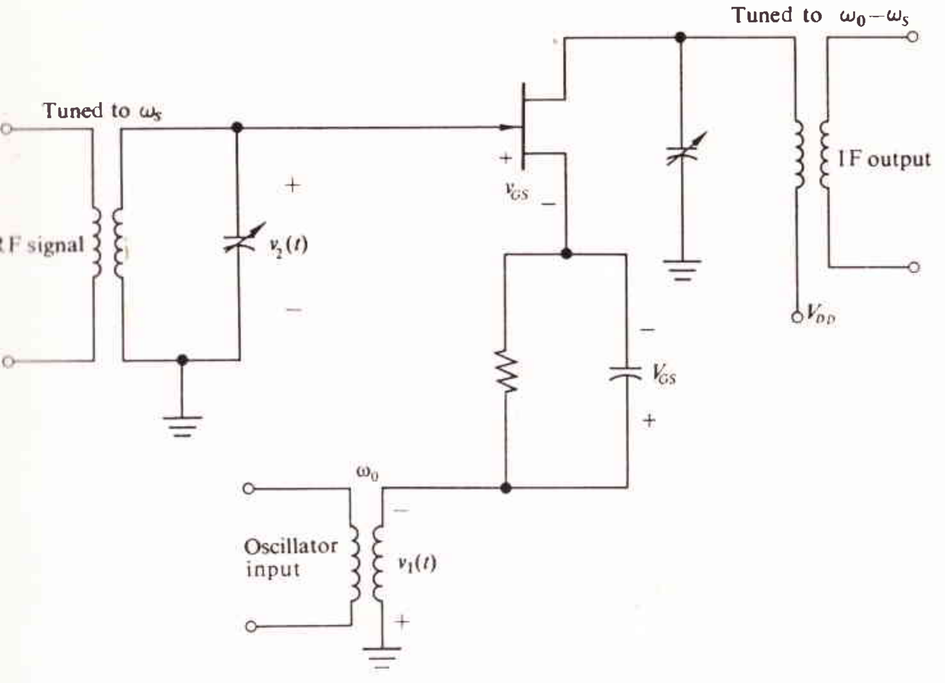


Fig. 7.2–1 (a) FET mixer.

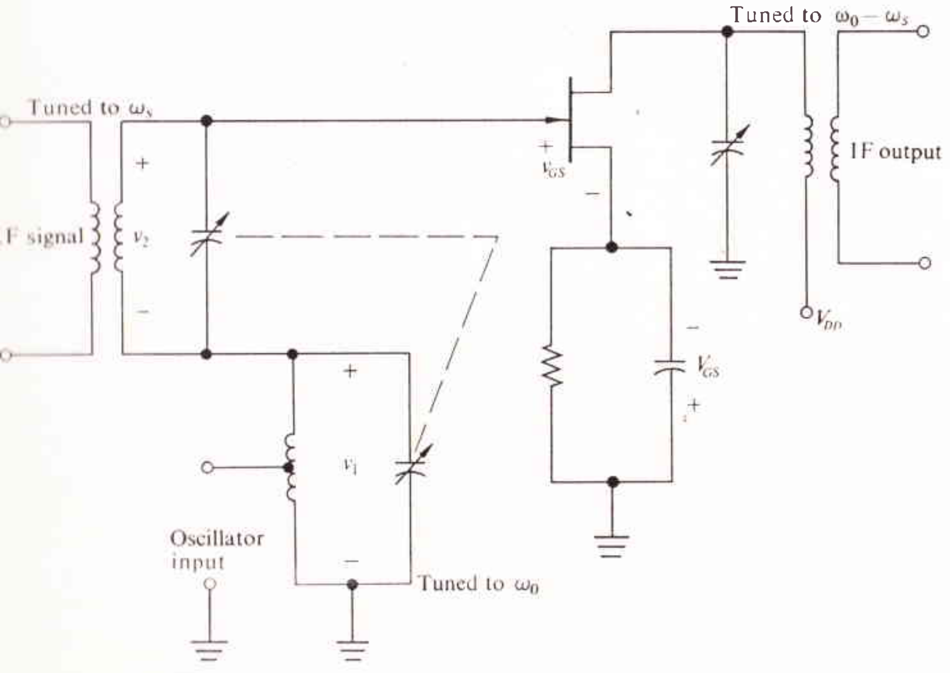


Fig. 7.2–1 (b) FET mixers (both signals to gate circuit).

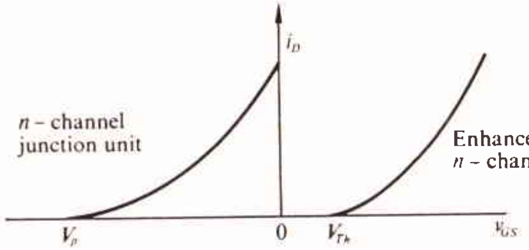


Fig. 7.2–2 FET characteristics.

If $v_2(t) = v_s(t)\cos\omega_s t$ while $v_1(t) = V_1\cos\omega_0 t$, and if we assume IF filtering at the output and consider only the frequency component of the output current in the vicinity of $\omega_0 - \omega_s$, we may define a large-signal conversion transconductance, G_c, as the envelope of the output current at the desired difference frequency divided by the envelope of the input signal voltage.

Equation (7.2–4) presents the conversion transconductance from ω_s to $\omega_0 - \omega_s$ as well as the ordinary large-signal amplifier transconductance:

$$G_c = \frac{\beta V_1}{2}, \qquad G_m = \beta(V_{GS} - V_{Th}). \tag{7.2–4}$$

Hence $G_c/G_m = V_1/2(V_{GS} - V_{Th})$. Now since $v_s(t) + V_1$ must be less than $V_{GS} - V_{Th}$ if one is not to exceed the "square law" region, the conversion transconductance can not exceed one-half the transconductance of the device when it is employed as an ordinary amplifier. Note that if V_1 is truly constant, then as long as the previous assumptions are maintained, G_c is independent of $v_s(t)$ and distortionless conversion results.

A similar expansion for the junction FET yields

$$G_c = \frac{I_{DSS}}{V_p^2}V_1, \qquad G_m = \frac{2I_{DSS}}{V_p^2}(V_p - V_{GS}). \tag{7.2–5}$$

If this device is biased midway between V_p and zero and if $V_1 \gg |v_s(t)|$, then the maximum obtainable G_c is one-quarter of the small-signal g_m evaluated at $V_{GS} = 0$ or one-half of the $G_m = g_m$ for the $V_{GS} = V_p/2$ bias point.

Bipolar Transistor Mixers

Figure 7.2–3 illustrates a possible bipolar transistor mixer.

When two time-varying voltages, $v_1(t)$ and $v_2(t)$, are applied across an ideal bipolar transistor junction, then

$$i_E(t) = I_{ES}e^{qV_{dc}/kT}e^{v_1(t)q/kT}e^{v_2(t)q/kT}. \tag{7.2–6}$$

If again $v_2(t) = v_s(t)\cos\omega_s t$ and $v_1(t) = V_1\cos\omega_0 t$ while $y(t) = v_s(t)q/kT$ and $x = V_1 q/kT$, then

$$i_E(t) = [I_{ES}e^{qV_{dc}/kT}]\left[I_0(x) + 2\sum_1^{\infty} I_n(x)\cos n\omega_0 t\right]\left[I_0(y) + 2\sum_1^{\infty} I_n(y)\cos n\omega_s t\right]. \tag{7.2–7}$$

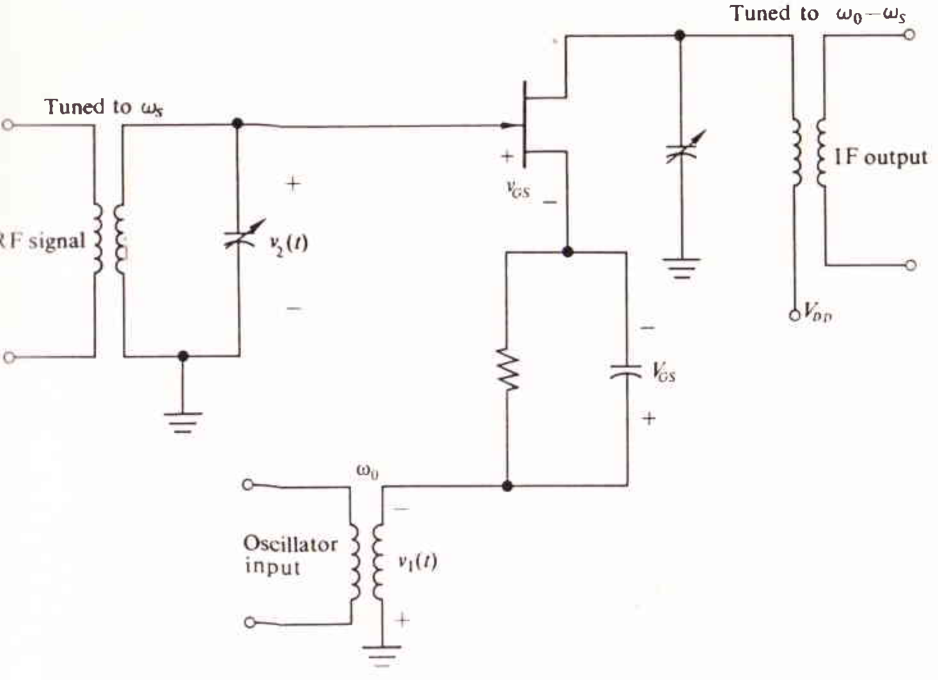


Fig. 7.2–1 (a) FET mixer.

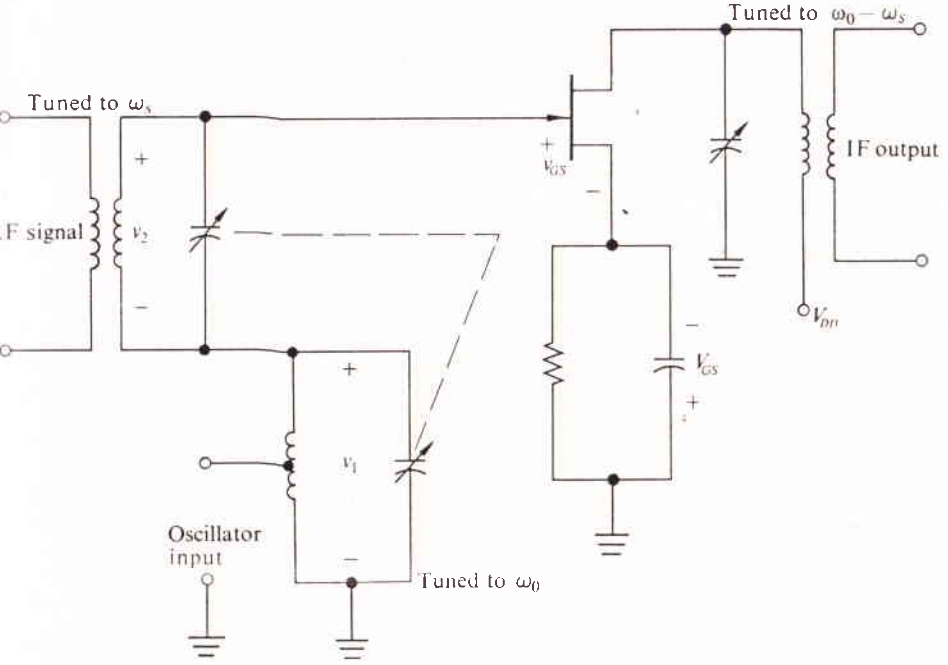


Fig. 7.2–1 (b) FET mixers (both signals to gate circuit).

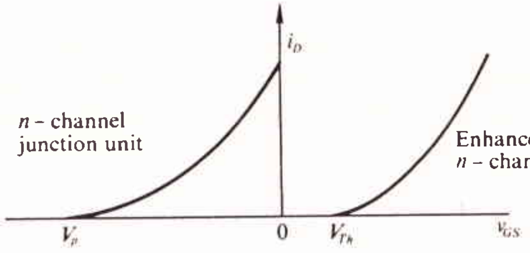


Fig. 7.2–2 FET characteristics.

If $v_2(t) = v_s(t)\cos\omega_s t$ while $v_1(t) = V_1\cos\omega_0 t$, and if we assume IF filtering at the output and consider only the frequency component of the output current in the vicinity of $\omega_0 - \omega_s$, we may define a large-signal conversion transconductance, G_c, as the envelope of the output current at the desired difference frequency divided by the envelope of the input signal voltage.

Equation (7.2–4) presents the conversion transconductance from ω_s to $\omega_0 - \omega_s$ as well as the ordinary large-signal amplifier transconductance:

$$G_c = \frac{\beta V_1}{2}, \qquad G_m = \beta(V_{GS} - V_{Th}). \tag{7.2–4}$$

Hence $G_c/G_m = V_1/2(V_{GS} - V_{Th})$. Now since $v_s(t) + V_1$ must be less than $V_{GS} - V_{Th}$ if one is not to exceed the "square law" region, the conversion transconductance can not exceed one-half the transconductance of the device when it is employed as an ordinary amplifier. Note that if V_1 is truly constant, then as long as the previous assumptions are maintained, G_c is independent of $v_s(t)$ and distortionless conversion results.

A similar expansion for the junction FET yields

$$G_c = \frac{I_{DSS}}{V_p^2}V_1, \qquad G_m = \frac{2I_{DSS}}{V_p^2}(V_p - V_{GS}). \tag{7.2–5}$$

If this device is biased midway between V_p and zero and if $V_1 \gg |v_s(t)|$, then the maximum obtainable G_c is one-quarter of the small-signal g_m evaluated at $V_{GS} = 0$ or one-half of the $G_m = g_m$ for the $V_{GS} = V_p/2$ bias point.

Bipolar Transistor Mixers

Figure 7.2–3 illustrates a possible bipolar transistor mixer.

When two time-varying voltages, $v_1(t)$ and $v_2(t)$, are applied across an ideal bipolar transistor junction, then

$$i_E(t) = I_{ES}e^{qV_{dc}/kT}e^{v_1(t)q/kT}e^{v_2(t)q/kT}. \tag{7.2–6}$$

If again $v_2(t) = v_s(t)\cos\omega_s t$ and $v_1(t) = V_1\cos\omega_0 t$ while $y(t) = v_s(t)q/kT$ and $x = V_1 q/kT$, then

$$i_E(t) = [I_{ES}e^{qV_{dc}/kT}]\left[I_0(x) + 2\sum_1^{\infty} I_n(x)\cos n\omega_0 t\right]\left[I_0(y) + 2\sum_1^{\infty} I_n(y)\cos n\omega_s t\right]. \tag{7.2–7}$$

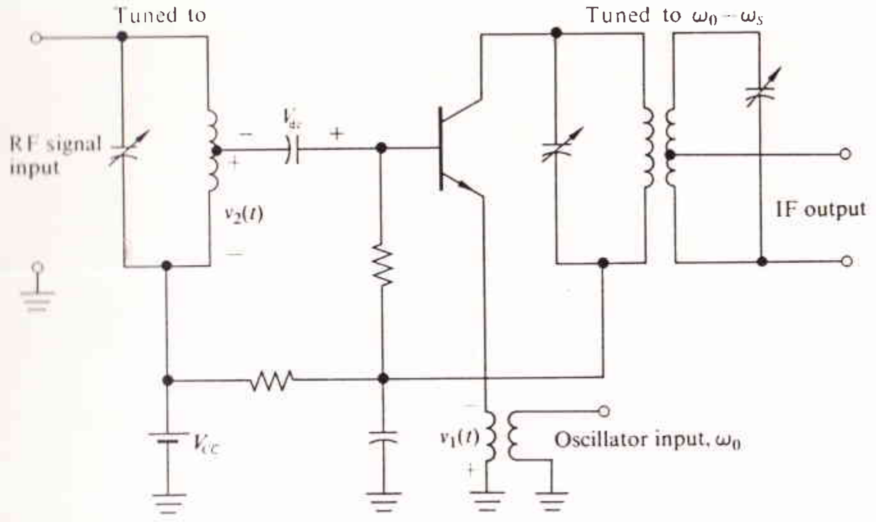


Fig. 7.2–3 Bipolar transistor mixer.

ve cross-multiply terms, we obtain

$$i_L(t) = [I_{ES}e^{qV_{dc}/kT}][I_0(y)I_0(x) + 2I_0(x)I_1(y)\cos\omega_s t + 2I_0(y)I_1(x)\cos\omega_0 t + 4I_1(x)I_1(y)[\cos\omega_s t\cos\omega_0 t + \cdots], \tag{7.2–8}$$

m which

$$I_{dc} = I_{ES}e^{qV_{dc}/kT}I_0(y)I_0(x), \tag{7.2–9}$$

$$I_{\omega_s} = 2I_{dc}\frac{I_1(y)}{I_0(y)}, \qquad I_{\omega_0} = 2I_{dc}\frac{I_1(x)}{I_0(x)}, \tag{7.2–10}$$

$$I_{\omega_0-\omega_s} = 2\frac{I_1(x)}{I_0(x)}\frac{I_1(y)}{I_0(y)}I_{dc}. \tag{7.2–11}$$

In order to have a "linear" mixer, one must have $I_{\omega_0-\omega_s}$ linearly proportional to nd in turn $v_s(t)$. If V_1 and hence x are assumed to be constant and if $V_1 \gg |v_s|$ that variations in v_s do not effect I_{dc}, then $I_{\omega_0-\omega_s}$ will vary with v_s as $I_1(y)/I_0(y)$ ies with v_s.

From the expansions for the modified Bessel functions with small arguments, may write

$$\frac{I_1(y)}{I_0(y)} \approx \frac{y}{2}\left(1 - \frac{y^2}{8} + \frac{y^4}{16}\right), \tag{7.2–12}$$

ich is within 2% for $y \leq 1$.

From Eq. (7.2–12) we see that for $I_1(y)/I_0(y)$, and hence $I_{\omega_0-\omega_s}$, to be linear with

---

respect to y to within 2% the condition $y^2/8 < 0.02$ or $y \leq 0.4$ and hence $|v_s| \leq$ 10.4 mV applies; for linearity to within 0.5%, the appropriate condition is $y \leq 0.2$ or $|v_s| \leq 5.2$ mV. Therefore, when $|v_s| \leq 10.4$ mV we may replace $I_1(y)/I_0(y)$ by $y/2$ and from Eq. (7.2–11) obtain

$$\frac{I_{\omega_0-\omega_s}}{V_s} = G_c = \frac{I_1(x)}{I_0(x)}g_{in}, \frac{I_{\omega_0}}{V_1} = G_m = 2\frac{I_1(x)}{xI_0(x)}g_{in}, \frac{I_{\omega_s}}{V_s} = G_s = g_{in} \tag{7.2–13}$$

$$G_c/G_m = x/2, \qquad G_c/G_s = I_1(x)/I_0(x), \tag{7.2–14}$$

where $g_{in} = qI_{dc}/kT$ and $v_s(t) = V_s$.†

Thus in this case G_c may never exceed G_s but may exceed G_m if $x > 2$. Since $I_1(x)/I_0(x)$ is within 70% of its asymptotic value of unity when $x = 2$ and within 86% of this value when $x = 4$, not much is to be gained from using values of x exceeding 6.‡

There are two useful ways to connect the signals to a differential-pair circuit to make a mixer. In one case the relatively large oscillator voltage is fed into one (or across both) of the differential-pair bases, while the relatively small signal voltage is fed across the emitter-base junction of the "constant current" transistor. Figure 7.P–3 illustrates a simplified version of such a circuit.

In the second case the oscillator voltage is used to control the "constant current" transistor, while the signal is fed into (or across) the differential-pair base circuits. Figure 7.4–4 provides an example of a balanced version of such a circuit.

As we shall see in Section 7.4, the most desirable way (from the distortion viewpoint) to operate a differential-pair mixer is so that both the oscillator and the "signal" operate the mixer linearly when they are considered as separate inputs. Under these circumstances, the mixing action takes place because the "gain" from the differential-pair base-base voltage to the output is directly proportional to the current from the "constant current" transistor's collector.

Overall system considerations will decide which type of drive is most suitable for a particular case. When the mixer generates its own oscillation (serves as a converter), then one would tend to confine the oscillator to the differential pair and feed the signal into the "constant current" transistor.

The results of Sections 4 6 and 4.8 are directly applicable to mixer calculations. For example, Table 4.6–2 and Fig. 4.6–4 indicate the limits on the drive if linearity is to be maintained, and provide the value of the fundamental term for the case in which the differential pair is driven directly by an oscillator signal that is large enough to

---

† Since the resultant conductances are independent of V_s, they may be calculated with a constant for $v_s(t)$ rather than the more complicated function of time.

‡ It is true that g_{in} may increase slowly with x since

$$I_{dc} = 1 + \frac{\ln I_0(x)}{qV_\lambda/kT}$$

(see Fig. 5.4–4, for example). For $x = 6$ and a 1 V or more dc drop across an emitter resistor, this increase will be less than 10%.

the (I_1/I_k) relationship to become nonlinear. From Table 4.6–2, we see that an ator drive of 78 mV between the bases leads to a fundamental term very close to value of I_k. Hence the conversion transconductance from a signal that influ- I_k to the output current of one side of the differential pair will be $\frac{1}{4}$ (another $\frac{1}{2}$ from the separation of the oscillator-signal product into sum and difference) of the small-signal transconductance (from base-emitter signal voltage to tor currents) of the "constant current" transistor.

he techniques of Section 4.8 may be used to linearize either the differential-pair eteristic or the characteristic of the "constant current" transistor.

alized Mixers

the FET, bipolar transistors, and differential pair account for most practical , it is useful to generalize the approach, both as a means of dealing with other s—vacuum tubes, for example—and as a means of examining the previous s when their characteristics depart from the usual shapes.

ne general approach is to consider the mixer as a gain-controlled amplifier in the local oscillator voltage controls the gain while the signal voltage is amplified. 7.2–4 illustrates one extreme case, where the gain has only two values and the local oscillator is biased so that the device spends half the time in each

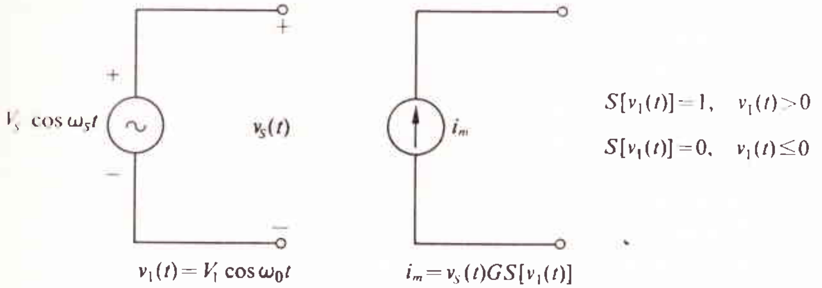

$V_s \cos\omega_s t$
+
−
$v_s(t)$
+
−
i_m
$S[v_1(t)]=1, \quad v_1(t)>0$
$S[v_1(t)]=0, \quad v_1(t)\le 0$
$v_1(t)=V_1\cos\omega_0 t$
$i_m=v_s(t)GS[v_1(t)]$

Fig. 7.2–4 On–off amplifier or switched amplifier as a mixer.

ice the output current multiplier is controlled by $v_1(t)$, it is periodic and can anded in a Fourier series with fundamental ω_0:

$$GS[v_1(t)] = \frac{G}{2} + \frac{2G}{\pi}\cos\omega_0 t - \frac{2G}{3\pi}\cos 3\omega_0 t + \cdots + .$$

ice the output current is the product of $v_2(t)$ and $GS[v_1(t)]$, it is apparent that

$$G_c = \frac{G}{\pi}.$$

nis technique of specifying the amplifier "gain" or transfer term with respect nput signal and then expanding the controlled "gain" in a Fourier series is a on one. Usually the gain or transfer function is a continuous function of the

oscillator voltage and it is evaluated by taking the partial derivative of i with respect to the signal voltage. (This procedure is based on the assumption that the signal term is small enough that it does not influence the gain. In any particular case the limits on v_s before intolerable distortion sets in must be determined.)

For example, if we differentiate the basic transistor junction equation, Eq. (7.2–6), with respect to v_2 and assume that v_2 approaches zero,

$$\left.\frac{\partial i_E}{\partial v_2}\right|_{v_2\to 0} = \frac{q}{kT} I_{ES} e^{qV_{dc}/kT} e^{v_1(t)q/kT}, \tag{7.2–15}$$

where

$$\left.\frac{\partial i_E}{\partial v_2}\right|_{v_2\to 0}$$

is the transfer conductance from $v_2(t)$ to the output current. The right side of Eq. (7.2–15) may be expanded as the first two brackets on the right side of Eq. (7.2–7). Again we find

$$G_c = \frac{I_1(x)}{I_0(x)}\frac{qI_{dc}}{kT} = \frac{I_1(x)}{I_0(x)} g_{in}.$$

In this example the initial approach is preferable, since an exact expression is possible for $I_{\omega_0-\omega_s}$ and hence the approximations are made at the last step rather than at the first step.†

As a second example consider the field-effect transistor for which the "gain" term is shown in Fig. 4.4–2. Since this term varies linearly as long as the oscillator voltage stays above V_p and below the junction turn-on point, the conversion transconductance is directly proportional to the oscillator voltage within this region. When the signal extends outside the linear region, then the sine-wave tip functions of Fig. 4.2–3 or Fig. 4.2–4 may be used to compute the fundamental term of the "gain." This term is directly proportional to the small-signal conversion transconductance.

Figure 7.2–5 illustrates the results for a peak clamping n-channel FET circuit similar to that used in Problem 7.5 (with the 100–Ω source-to-ground resistance set to zero).

To account for the drive up to +0.7 V, I_{DSS} is replaced by $I^*_{DSS} = I_{DSS}(1 - 0.7/V_p)^2$ and V_p is replaced by $V^*_p = |V_p| + 0.7$, and clamping is now presumed to occur at zero volts on the shifted curve. (As discussed in Chapter 5, the actual turn-on voltage is a function of the gate-ground resistor. When this resistor is several megohms, the turn-on voltage may be only about 0.3 V.)

As we shall see when (V_1/V^*_p) exceeds 0.5, many of the desirable low-distortion properties of the FET mixer are lost. One possible argument for operation above this point is that the dc current is reduced, which leads to less dissipation and usually less device noise.

---

† Yet another approach to this problem would be simply to expand $e^{qv_2(t)/kT}$ and retain only the first several terms, and then to expand the remaining expression as in the first two terms of Eq. (7.2–7) and take the resultant product.

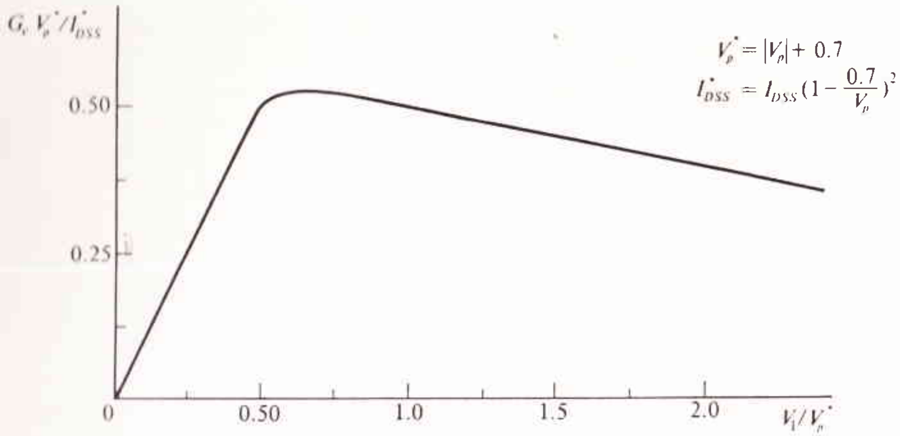


Figure 7.2–5

While we have done the two examples, the usual difficulty with this approach is evaluating the Fourier coefficients of the "gain" function. In the general case this nnot be done in a simple analytic manner; it must be done graphically, by a comter analysis, or experimentally. Of course, if the problem is attacked experimentally e will logically measure the conversion transconductance directly instead of fiddling ound with intermediate steps.

SERIES RESISTANCE IN MIXERS

order to operate, a mixer must have a nonlinearity. As we might expect (see Section , for example), adding a resistor in series with the emitter of a bipolar transistor, th the source of an FET, or with the emitters of a differential pair always tends to earize the characteristic. Therefore, such a resistor will always reduce the conrsion transconductance (if a constant local oscillator voltage is assumed). Since real devices have at least a minimum amount of inherent series resistance, to say thing of generator output impedance, one will find that the calculations of the evious section provide an upper limit on the expected performance of real mixers.

The effect of a series resistance in modifying the incremental slope of the i_1 vs. v_1 aracteristic is clearly indicated by Fig. 4.8–2. It is a straightforward matter to solve $\partial i_1/\partial v_1$ from the equations of Section 4.8:

$$\left.\frac{\partial i_1}{\partial v_1}\right|_{V_1 \cos \omega_0 t + V_{dc}} = \frac{i_1}{kT/q + i_1 R} = \frac{g_{in} i_1}{I_{dc} + g_{in} R i_1}, \qquad (7.3\text{–}1)$$

ere $g_{in} = qI_{dc}/kT$ and i_1 is the current which flows with $v_1 = V_1 \cos \omega_0 t + V_{dc}$. When $g_{in}R \gg 1$, Eq. (7.3–1) reduces to $1/R$ for $v_1 - V_0 > 0$ and to zero for $- V_0 \leq 0$. For $g_{in}R \to 0$, Eq. (7.3–1) reduces to

$$\frac{q i_1}{kT} = \frac{q}{kT} I_{ES} e^{qV_{dc}/kT} e^{qV_1 \cos \omega_0 t/kT},$$

which is of course identical with Eq. (7.2–15). For intermediate values of $g_{in}R$, no simplified form for the gain function is possible.

A possible, and relatively simple, approach is first to find a reasonable relationship between G_c and G_m for the extreme cases of zero resistance and dominant resistance, and then, using Fig. 4.8–6, to determine G_m vs. V_1 and extrapolate to approximate G_c vs. V_1. The results will not be exact, however, the series resistance is rarely known exactly either. The best that one can do here is to get a feel for what is going on. If "exact" results are desired, then measurements are so straightforward that pages of analytic manipulation are completely unjustified.

When the junction is dominated by R, the slope of the overall characteristic is $1/R$ when $v_1 > V_0 = (kT/q) \ln (I_{dc}/I_{ES})$ (see Fig. 4.8–3) and is zero otherwise. This means that the "gain" function is a rectangular wave of period $T = 2\pi/\omega_0$, as shown in Fig. 7.3–1. If I_{dc} is assumed to be constant, then the conduction angle $\theta_c = 2\phi$ is related to $V_1/I_{dc}R$ by Table 4.8–1. The conduction angle then tells us the relative nonzero width of the "gain" function from which the fundamental component of the Fourier series may be computed. The conversion transconductance is then one-half of this fundamental coefficient. Table 7.3–1 indicates representative values.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACIO
Y BIBLIOTECA
MONTEVIDEO - U

Table 7.3–1

$\theta_c = 2\phi$	$V_1/I_{dc}R$	G_cR	G_c/G_m
360°	1.00	0.00	0.00
300°	1.063	0.167	0.17
270°	1.333	0.217	0.24
210°	1.667	0.29	0.36
180°	3.1416	0.317	0.63
150°	10.0	0.305	1.7
90°	20.5	0.23	2.4

Figure 7.3–2 plots G_c/G_m vs. $V_1/I_{dc}(r_{in} + R)$ for the two cases $R = 0$ and $R \gg 1/g_{in}$ [the $R = 0$ expression comes from Eq. (7.2–14)].

From this figure it is hypothesized that for values of $g_{in}R$ near unity it will be plausible to relate G_c/G_m to the normalized value of $V_1/I_{dc}(r_{in} + R)$ by a factor of $\frac{1}{3}$. For values of $g_{in}R$ between 0 and 1 this factor should lie between $\frac{1}{2}$ and $\frac{1}{3}$; for values

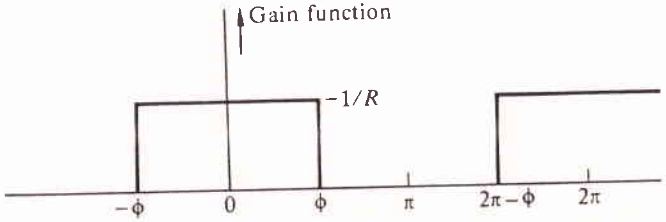


Fig. 7.3–1 G_c/G_m for a bipolar transistor mixer with added series resistance.

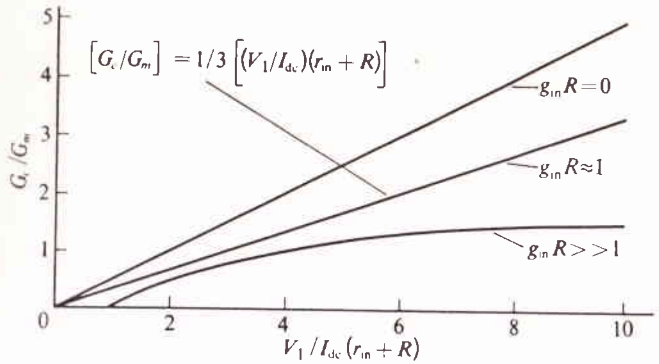


Fig. 7.3–2 Data for Eq. (7.3–2).

etween 1 and infinity it should lie between $\frac{1}{3}$ and $\frac{1}{6}$. Normally one would expect to ninimize the resistance, and hence the $g_{in}R < 1$ region would be expected to be the nteresting one.

For example, if $I_{dc} = 1$ mA, $R = 26\,\Omega$, and $V_1 = 6 \times 52$ mV $= 312$ mV, then rom Fig. 7.3–2 $G_c \approx 2G_m$, while from Fig. 4.8–6 $G_m = 0.295(52\,\Omega) = 5700\,\mu$mho. Vith $R = 0$, $I_{dc} = 1$ mA, and half the oscillator drive (that is, $V_1 = 156$ mV), then $= 6$ and $G_c = 34{,}500\,\mu$mho, or approximately three times as much as in the previous ase (a glance at Fig. 4.8–5 will indicate that maintaining the 312 mV drive would ot increase the G_c appreciably for this second case).

A similar approach is possible for FET's and differential pairs. In either case, hen R times the quiescent small-signal transconductance becomes appreciable, the onversion transconductance will fall sharply.

Vogel† has considered the effect of serious source resistance on the peak-clamping ET of pages 301 and 302 (Fig. 7.2–5). When certain of Vogel's results are simplified nd translated into the peak clamping terms V_p^* and I_{DSS}^* of Fig. 7.2–5, then the $_cV_p^*)/I_{DSS}^*$ vs V_1/V_p^* curve of Fig. 7.2–5 is found to be depressed by the addition of a eries source resistance, R_s.

As a reasonable first approximation to the effect of the series resistance, one takes e value of $(G_cV_p^*/I_{DSS}^*)$ from Fig. 7.2–5, for the particular value of V_1/V_p^* of interest, nd divides it by $(1 + q)^3$, where $q = (2I_{DSS}^*R_sz)/V_p^*$ and z is a term that accounts for e shift in the dc bias current with increasing drive; z is plotted vs. V_1/V_p^* in Fig. 3–3.

Suppose that $(I_{DSS}^*/V_p^*) = 2000\,\mu$mho, that $(V_1/V_p^*) = 0.5$, and that R_s is 20 ohms. this case z will be 0.5, $q = 0.04$, and $(1 + q)^3$ will be 1.125. Therefore G_c will be duced from $1000\,\mu$mho for $R_s = 0$ to $889\,\mu$mho for $R_s = 20$ ohms. [From Fig. 2–5, $(G_cV_p^*/I_{DSS}^*) = 0.5$ when $(V_1/V_p^*) = 0.5$.]

S. Vogel, "Nonlinear Distortion and Mixing Process in Field Effect Transistors." *Proc. EE*, 55, No. 12, pp. 2109–2116 (Dec. 1967).

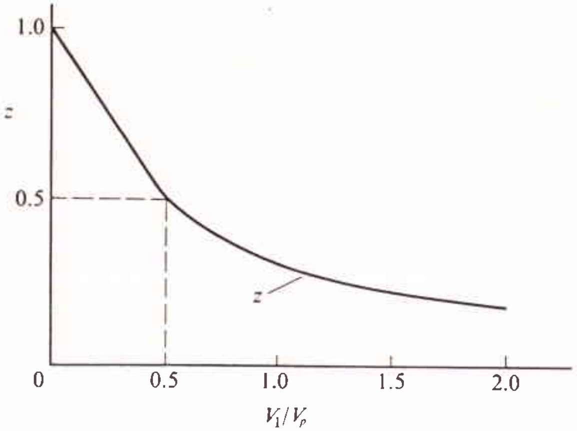


Figure 7.3–3

Doubling the drive voltage so that $(V_1/V_p^*) = 1$ will reduce z to approximately 0.32, q to 0.0256, and $(1 + q)^3$ to 1.079. In this case $G_cV_p^*/I_{DSS}^*$ is again 0.5; hence now G_c is $927\,\mu$mho instead of $1000\,\mu$mho.

If R_s is made much larger, say 200 ohms, then for the $(V_1/V_p^*) = 0.5$ case, one will find that $z = 0.5$, $q = 0.4$, and $(1 + q)^3 = 2.74$, so that now G_c is reduced from 1000 μmho for $R_s = 0$ down to only $365\,\mu$mho.

7.4 PRACTICAL MIXER CIRCUITS

In this section we shall discuss the interfering-signal problem encountered in bipolar transistor, differential-pair, and FET mixers. In the next section we will indicate how to combine the mixing and oscillating functions to produce "converters" from each of these devices.

Before we proceed with particular circuits it is desirable to reconsider the problems that are likely to be encountered, both so that we can avoid them or combat them and so that we can compare different circuits with respect to them.

Mixer Problems

1) Output signals at the IF frequency arising from other than the desired input signal.
2) Distortion of the modulation of the desired input signal.
3) Transmission of the local oscillator frequency to the input circuit. (If the local oscillator signal reaches the antenna, it may be radiated and serve as an interfering signal to other receivers.)
4) Noise generated in the mixer stage.
5) Inadequate gain in the mixer stage.

Figure 7.4–1 illustrates a frequency spectrum showing a number of possible signals that may cause unwanted components at the mixer output's IF frequency.

a) If a signal at $\omega_{\text{image}} = \omega_0 + \omega_{\text{IF}}$ reaches the mixer, the difference frequency will be ω_{IF}. The only remedy is adequate filtering in front of the mixer. In a new

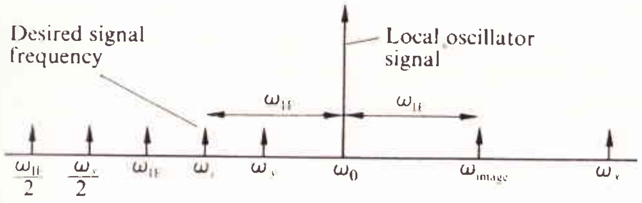


Fig. 7.4–1 Desired and possible interference signals in a superheterodyne receiver.

stem the choice of the highest possible IF frequency will ease the image-rejection oblem.

A signal at $\omega_s/2$ will cause trouble if the RF stage has enough distortion to proce a second-harmonic term from the signal or if the mixer provides a beat frequency tween the ω_0-term and the second harmonic of the incoming signal. From Eq. 2–8) for the bipolar transistor,

$$I_{\omega_0-\omega_s/n} = 2\frac{I_1(x)}{I_0(x)}\frac{I_n(z)}{I_0(z)}I_{dc},$$

ere z is the normalized envelope of the (ω_s/n)-term. When z is small, $I_n(z)/I_0(z)$ y be approximated as $z^n/2^n n!$, so that

$$\frac{G_c(\omega_s)}{G_c(\omega_s/n)} \approx \frac{2^{n-1}n!}{z^{n-1}}.$$

equate filtering to remove the half-frequency terms must be provided in front of T or square-law RF amplifiers; otherwise, these devices will double the input quency and no later stage will remove it. With "linear" RF amplifiers all premixer ering is effective in reducing subharmonic terms.

The signal ω_x is at $\omega_0 + \omega_s$. It will cause trouble in any system where a beat quency with $2\omega_0$ is possible. Hence it will cause trouble in transistor mixers but in true square-law FET circuits or in differential-pair circuits, provided that oscillator driving voltage is free of second harmonics.

The signal ω_y falls at $\omega_{image}/2$; hence an FET RF amplifier will shift it to ω_{image} l, unless the filtering *between* this RF amplifier and the mixer removes this distor- term, there will be an unwanted output term. A single-ended bipolar transistor er will produce an output similar to the $\omega_s/2$ case.

Obviously there are many other potential sources of interference; however, above terms suffice to illustrate the problem. As a numerical example, consider se where

$$\omega_{IF} = 455\,\text{kHz}, \qquad \omega_s = 600\,\text{kHz}, \qquad \omega_0 = 1055\,\text{kHz}$$

hence

$$\omega_{s/2} = 300\,\text{kHz}, \qquad \omega_y = 755\,\text{kHz}, \qquad \omega_{image} = 1510\,\text{kHz},$$

$$\text{and} \qquad \omega_x = 1655\,\text{kHz}.$$

A single-tuned circuit in front of a simple mixer that is supposed to handle 5 kHz modulation should not have a Q much above 8 at 600 kHz; if it does, tracking problems will become difficult and overall system sideband cutting will become excessive. Such a tuned circuit will reduce ω_y by a factor of only 4 with respect to the transmission of ω_s. Since the signal at ω_y may be four or more times as big as ω_s to start with the RF tuned-circuit outputs may be equal. If so, and if both signals have amplitudes of 10 mV, then in a bipolar transistor mixer's output the unwanted term will be 25% of the desired signal term. In an FET mixer or a differential pair, the main distortion term from a 755 kHz input would be at 300 kHz and would be rejected by the IF filter.

To summarize: An FET or differential-pair mixer should be superior to a bipolar mixer from the viewpoint of unwanted signals. At least two stages of filtering are desirable in front of the mixer. A linear, automatically gain-controlled RF amplifier is desirable in front of the mixer. The gain reduces the importance of the mixer noise, while the gain control allows the circuit to keep excessive signals from reaching the mixer. Excessive signals inevitably lead to increased distortion. In spite of our outline of the ideal situation, many simple receivers have the mixer as the first stage and employ a bipolar transistor as the active device in the circuit. Figure 7.4–2 illustrates such a circuit.

It is presumed that the impedance from the base to ground is small at both ω_0 and ω_{IF} and that the impedance from emitter to ground is small at ω_s and ω_{IF}. Failure to satisfy these assumptions leads to reduced conversion transconductance as well as the possibility of oscillation at the IF frequency. If oscillation occurs, it should be

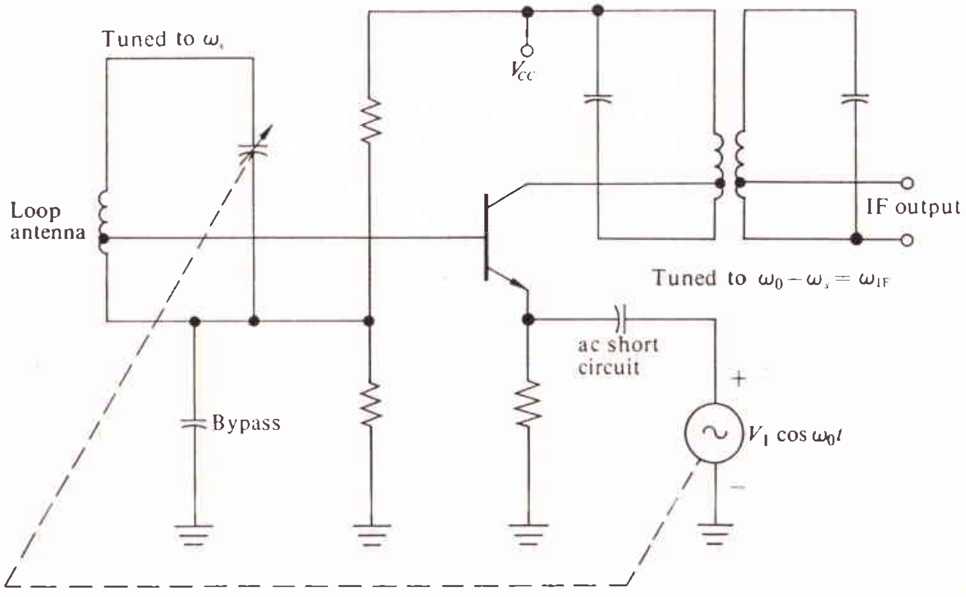


Fig. 7.4–2 Bipolar transistor mixer.

rable by a reduction of the base impedance at ω_{IF}. If necessary, a series trap tuned ω_{IF} could be connected from the base to ground.

Both input and output transformers are tapped down not only for the previous asons, but also so that the transistor impedances will not cause excessive detuning the circuits.

As has been pointed out, it is highly desirable to prevent large signals of any kind xcept those from the local oscillator) from reaching the mixer input. Hence it is ot sufficient to control the mixer-stage gain or conversion transconductance, say, controlling I_{EQ}. One should control an attenuator in the input circuit also. igure 7.4-3 illustrate the key portions of the addition of such a circuit to Fig. 7.4-2.

From the results of Chapter 4 and Chapter 5 we may write the impedance of the ode to the signal frequency as

$$G_d = \frac{I_k q}{kT} \frac{2I_1(x)}{xI_0(x)}. \tag{7.4-1}$$

we keep the input signals at or below 26 mV (as from Section 7.2 we know we ould), then $2I_1(x)/xI_0(x)$ is nearly unity and G_d is directly proportional to I_k. Hence creasing I_k will shunt the input and reduce the mixer driving signal. Increasing I_k ill also reduce I_{EQ} of the mixer and hence will reduce G_c, yielding further AVC tion. The driving source for I_k will be derived from one of the amplitude detector rcuits of Chapter 10. This effective diode shunting conductance reduces the Q of the input-tuned circuits; however, it only does so when the IF output is large and hence when the desired signal is present, one hopes, at a reasonable strength.

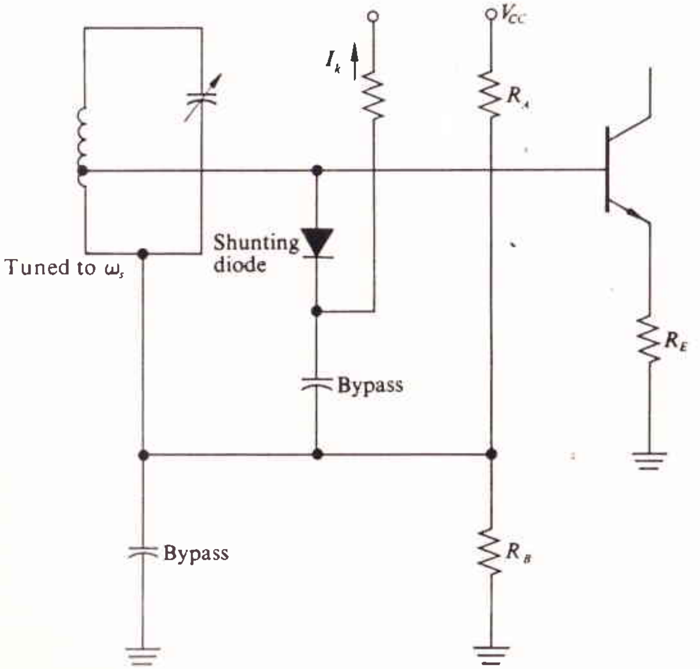


Fig. 7.4-3 Input AVC for mixer.

Differential-Pair Mixers

Figure 7.4-4 shows a possible circuit.

If Q_2 and Q_3 are truly identical and if the input and output transformers are truly balanced, then no oscillator voltage will appear across either the input or the output transformer. Removal of the oscillator voltage from the input removes the

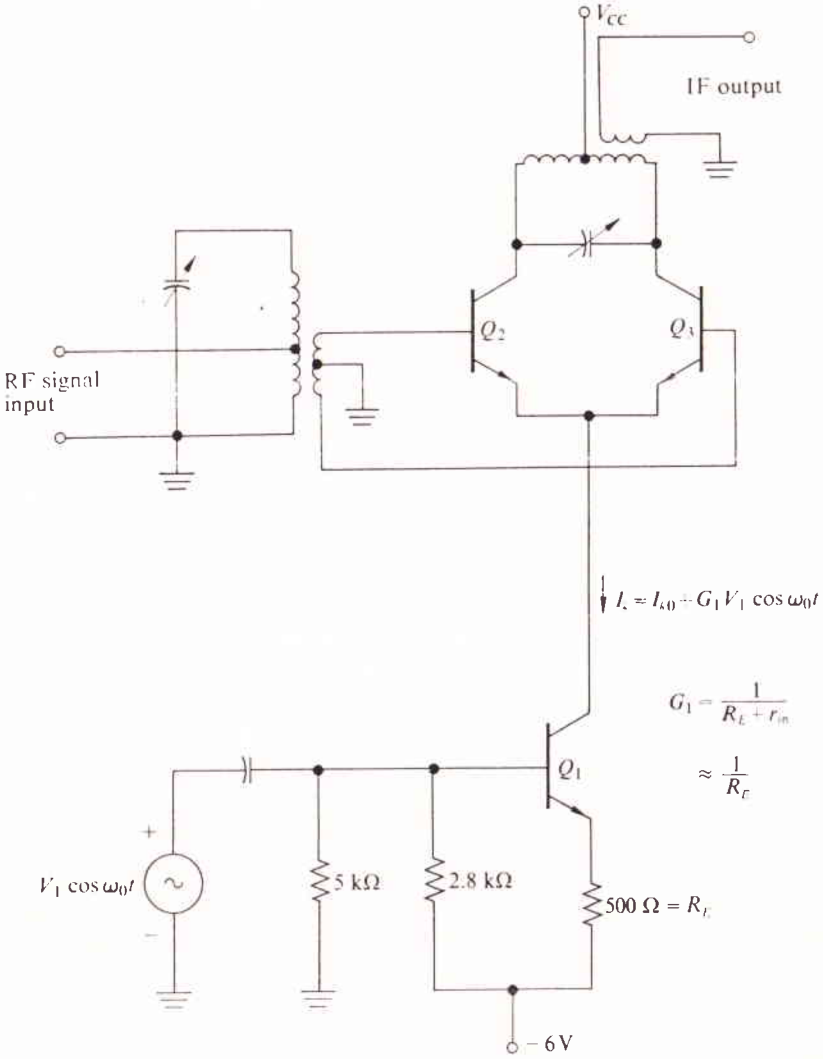


Fig. 7.4-4 Balanced mixer with differential pair and "constant current" driver.

adiation problem; in addition, removing the large oscillator current component om the output reduces the strain on the IF transformer. (In the single-ended stage, first IF transformer must reduce the large oscillator component sufficiently that does not cause nonlinear operation of the first IF amplifier.)

In addition, because of the symmetrical characteristics of the differential pair, cond-harmonic cross products should not be generated by this circuit. To avoid en harmonic terms completely, the oscillator voltage must be a true sine wave and variations in the collector current of Q_1 must be kept small enough that no oscil-or harmonics are generated. Measured data are quoted by RCA for such a ferential-pair circuit in which, for $V_1 = 141$ mV and an untuned input (so that no erfering signal reduction from input filtering is present), only three interference ms (except the image frequency and a component at the IF itself) are within 70 dB the desired signal. The first of these interference terms corresponds to ω_y in Fig. 1–1 ($2\omega_y - \omega_0 = \omega_{IF}$), the second two to ω_x and its twin ($2\omega_0 - \omega_x = \omega_{IF}$, $- 2\omega_0 = \omega_{IF}$). For $I_{k0} = 2.5$ mA all of these three terms were reduced by a factor $\frac{1}{60}$ or more below the level of the desired output term. The removal of the ω_y erference term requires exact symmetry in the differential-pair characteristic, ereas the removal of the ω_x-terms requires an absolutely pure oscillator voltage. either of these ideals will ever be achieved in practice. The ω_y-term is the most oublesome, since it is the closest one to the desired signal and hence the most ficult to filter out initially. The 60/1 reduction in the ω_y-term that is obtained by ing a practical differential pair rather than an ideal single-ended bipolar stage is pressive.

T Mixers

gure 7.4–5 shows another FET mixer circuit. In this case the high-impedance te can normally be connected directly across the input-tuned circuit. The oscillator ltage feedback will occur via the gate-source capacitance, and hence will be mini-ized if the minimum value of C_t is much greater than C_{GS}. If the device is biased that the peaks of the oscillator plus signal voltages never swing it out of the square-v region and if the oscillator voltage is a pure sine wave, then the only theoretical erference term (besides the image frequency and the IF frequency) will be at half IF frequency. Again, if the second harmonic is present in the oscillator voltage, en the ω_x- and ω_x'-terms will come through with relative values proportional to amount of this second harmonic. So long as the oscillator voltage is restricted om swinging to $V_p/2$ (this sacrifices some conversion transconductance), an increased nal level does not lead to distortion in the FET case; hence input AGC (or AVC) is t necessary. Variation of the conversion transconductance is achieved by varying oscillator voltage (see Eq. 7.2–4 or Eq. 7.2–5).

Many other mixer possibilities exist, but they will not be examined in detail. ost of them, be they multigrid vacuum tubes, beam-switching vacuum tubes, or al-gate MOS transistors, can be analyzed in a straightforward fashion, using the in-controlled amplifier approach of Section 7.2.

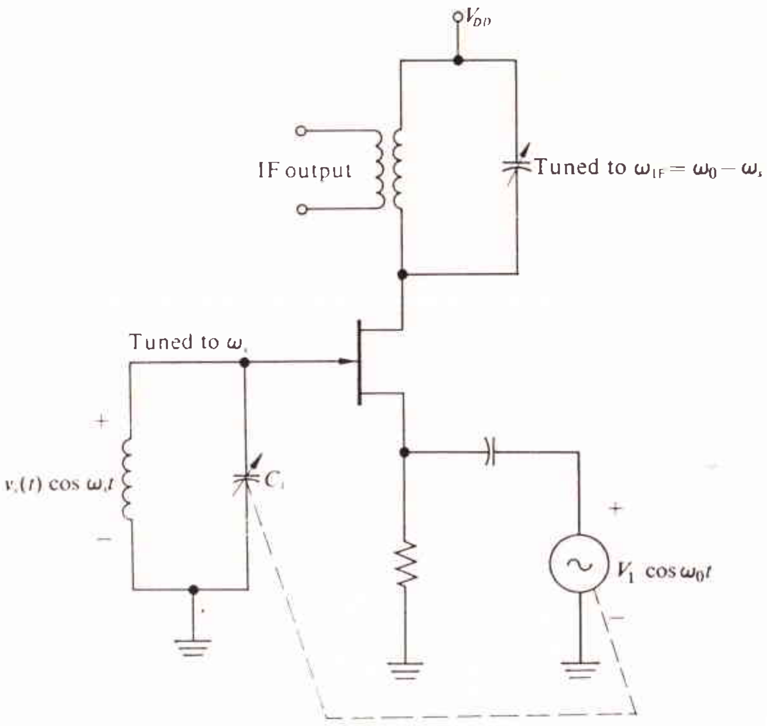


Fig. 7.4–5 FET mixer.

7.5 SEMICONDUCTOR CONVERTER CIRCUITS

Though a separate oscillator and a separate mixer circuit can normally be designed so that each does its own job best, it is possible to combine the two functions in a single active device. This combination is known as a converter. High-quality receivers usually keep these two functions separate, whereas most mass-produced receivers combine them.

Figure 7.5–1 shows a typical bipolar transistor converter stage. (A two-power-supply version is shown to reduce the complexity of the drawing slightly. The reader should have no difficulty in visualizing this circuit in a single-power-supply form.)

In such a circuit we design the oscillator circuit to give the driving level the value of $x = qV_1/kT$ desired for the mixer operation. In designing the oscillator we would like to be able to neglect the input signal circuit and the output IF circuit. Normally one does neglect them, and then checks up by showing that the voltage drops across these impedances caused by the calculated currents are truly negligible.

As in all mixing situations, one must worry about oscillator amplitude variation across the band, signal and oscillator circuit tracking, and mixer interference and distortion problems. As usual, compromises will be necessary. For example, the amplitude stability of the circuit may be improved by increasing the amplitude of

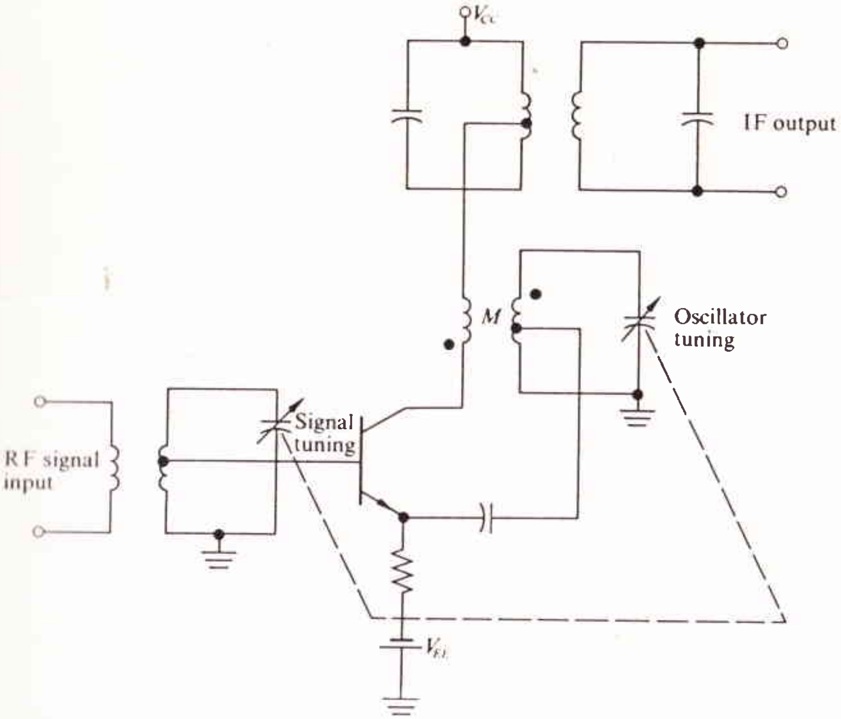


Fig. 7.5–1 Bipolar transistor converter.

oscillations, while the interfering output terms from oscillator harmonics or the culties with excessive oscillator voltage in either the IF or the input circuit are all imized by decreasing the amplitudes of the oscillations to their smallest possible ies.

Any of the previous mixer circuits may be turned into a converter by combining ith the appropriate oscillator circuit from Chapter 6.

Example 7.5–1 For the converter shown in Fig. 7.5–2, find an expression for

tion. If at the oscillator frequency $\omega_0 = 1/\sqrt{L_3C_3} = 1.5 \times 10^7$ rad/sec the edance of the input-tuned circuit is an effective short circuit compared with the -emitter impedance of the transistor, then the oscillation amplitude and frequency he converter may be found by grounding the base of the transistor and employing esults of Section 6.4. Specifically,

$$\omega_0 = 1.5 \times 10^7 \text{ rad/sec},$$

$$\frac{G_m(x)}{g_{mQ}} \approx \frac{G_L}{n g_{mQ}} = \frac{G_L L_3}{g_{mQ} M_{34}} = 0.19,$$

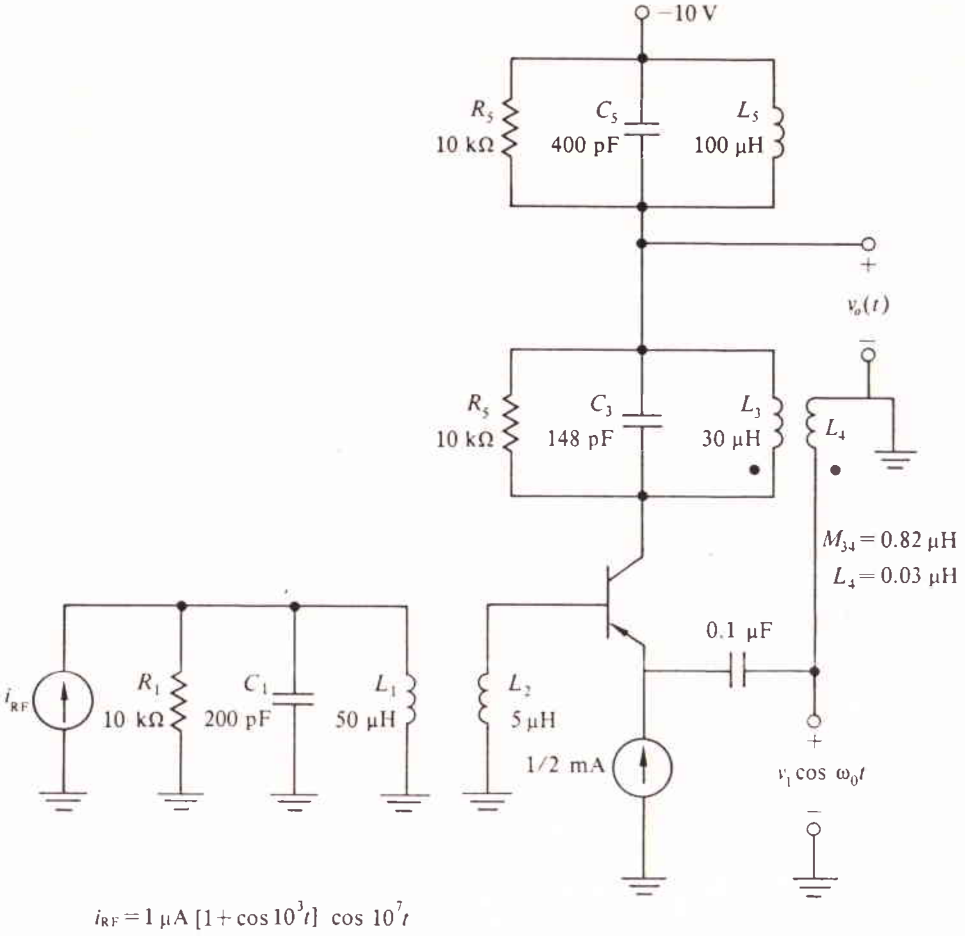


Figure 7.5–2

and from Fig. 4.5–6 $x = 10(V_1 = 260$ mV). In addition, since the loop gain is much less than unity at any other frequency, spurious oscillations at the RF or IF frequencies are not possible.

If we also assume that the input RF-tuned circuit is not loaded by the transistor, then the transistor base voltage is given by

$$v_B(t) = \frac{i_{RF} R_1 M_{12}}{L_1} = 1.57 \text{ mV}(1 + \cos 10^3 t) \cos 10^7 t,$$

since the bandwidth of the input-tuned circuit ($BW_{RF} = 1/R_1C_1 = 5 \times 10^5$ rad/sec) is sufficient to pass i_{RF} undistorted.

Now with the aid of Eq. (7.2–13), we obtain the IF component of collector rent in the form

$$i_{IF}(t) = -\alpha v_B(t)G_c = -\alpha v_B(t)g_{in}I_1(x)/I_0(x).$$

s IF component of collector current is extracted by the output-tuned circuit = $1/\sqrt{L_5C_5} = 5 \times 10^6$ rad/sec, $\mathrm{BW_{IF}} = 2.5 \times 10^6$ rad/sec) to yield

$$v_o(t) = (-10\ \mathrm{V}) - R_5 i_{\mathrm{IF}}(t),$$

ich with $x = 10$ reduces to

$$v_o(t) = (-10\ \mathrm{V}) - (286\ \mathrm{mV})(1 + \cos 10^3 t)\cos 5 \times 10^6 t.$$

To check the assumptions made in obtaining $v_o(t)$, we first obtain $|Z_B(j\omega_0)|$, ere $Z_B(p)$ is the impedance of the input circuit "seen" at the base of the transistor. ω_0, R_1 may be neglected compared with L_1 and C_1 to yield

$$|Z_B(j\omega_0)| = \left|\omega_0 L_2(1 - k_{12}^2) - \omega_0\left(\frac{L_2}{M_{12}}\right)^2\left[\frac{L_1}{(\omega_0/\omega_{\mathrm{RF}})^2 - 1}\right]\right|$$

$$\approx 48\ \Omega,$$

ich is indeed negligible compared with $(1 + \beta)/g_{in} = 5.2\ \mathrm{k\Omega}$, which is the transistor pedance at its base terminal. Hence our calculations of the oscillation frequency l amplitude neglecting the input circuit is justified.

With the techniques of Section 5.5, the equivalent "linear loading" of the tranor on the input-tuned circuit is readily shown to be

$$G_{NL} = \frac{I_{\mathrm{RF}}}{V_{\mathrm{RF}}} = g_{in}/(1 + \beta);$$

L may be reflected to the input of the RF-tuned circuit as $G_{NL}(M_{12}/L_1)^2 =$ 08 kΩ; hence as a first approximation the transistor loading may be neglected. tually the transistor in this example does decrease the impedance of the inputned circuit by 5%, and thus the IF component of $v_o(t)$ will be 5% below the value eviously calculated.

TUNED NARROWBAND SMALL-SIGNAL AMPLIFIERS

e usual receiver requires selectivity and gain both before and after the mixer or nverter circuit.

The sections in front of the mixer are known as RF stages. These stages should linear, to prevent the generation of cross modulation or the generation of distortion oducts that will interact in the mixer to cause signals in the IF band. In the freency bands where receiver noise is a limiting system parameter, these stages should ve adequate gain to prevent the higher noise of the mixer stage from limiting the ceiver's performance. In addition, these RF stages should provide the desired ectivity to prevent undesired signals from reaching the mixer.

The fixed tuned stages between the mixer and the demodulator circuit are known the IF stages. In AM systems these IF stages must be linear, while in FM systems

---

their amplitude characteristics may be markedly nonlinear. The IF stages should have high selectivity without causing distortion in the desired passband. In general, these fixed tuned stages provide the bulk of the gain between the input and the demodulator.

Thus the RF stages are normally low-gain, relatively broadband tunable stages, while the IF stages are high-gain, relatively narrowband, fixed tuned stages. If one keeps these differences in mind, the same general techniques may be utilized for designing both circuits.

Since one of the purposes of these stages is selectivity, the design of their selective networks is important. However, such design is not the main focus of this book.

Figure 7.6–1 illustrates the selectivity obtainable with a commercial seven-pole mechanical filter,† and Fig. 7.6–2 illustrates a possible driving and isolating circuit for a 455 kHz version of such a filter. A great variety of such filters is available, some at relatively low prices.

The concentration of selectivity in one "box" and broadband gain in a different "box" is one reasonable way to design a receiver. The field adjustments required to place seven separate poles in a Butterworth or Tschebyscheff arrangement are not desirable under any circumstances; hence, even if the filter were to be built up of lumped elements, one might well choose to have the filter "computer designed" and placed in the circuit as a unit.

What we will point out here is the basic approach to designing narrowband small-signal amplifiers that work into single- or double-tuned networks and are driven from single- or double-tuned circuits.‡

We shall use the narrowband y-model for the active circuit to provide an analysis that will be valid for single vacuum tubes, FET's, or bipolar transistors, as well as for differential pairs, cascodes, or other discrete or integrated circuit combinations.

Figure 7.6–3 shows the basic arrangement for a single stage. The y-parameters are defined at a single Q-point by the equations

$$I_1(p) = y_{11}(p)V_1(p) + y_{12}(p)V_2(p) \tag{7.6–1}$$

and

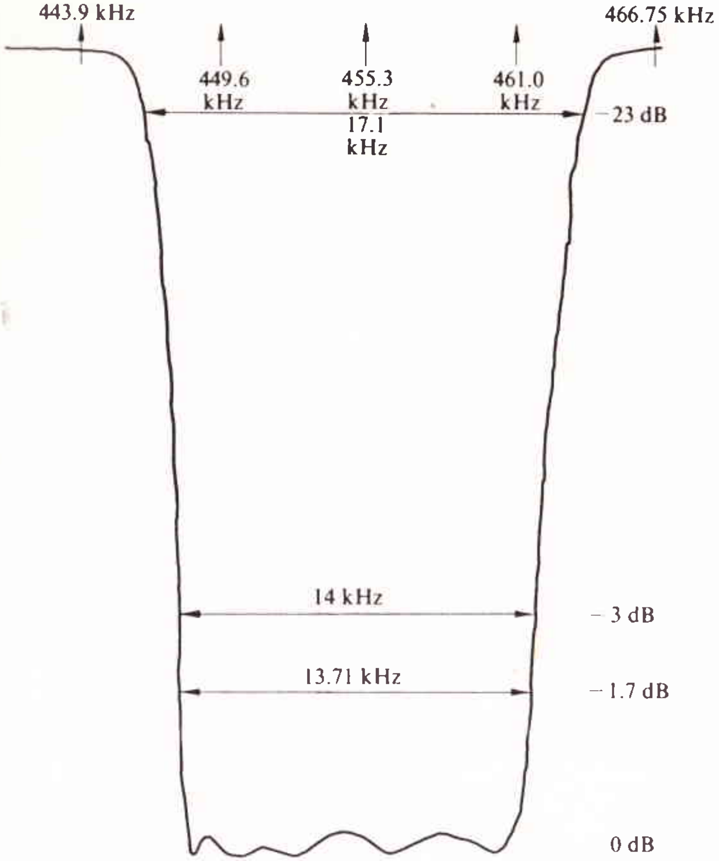
$$I_2(p) = y_{21}(p)V_1(p) + y_{22}(p)V_2(p). \tag{7.6–2}$$

Although these parameters may be calculated from some other device model, they

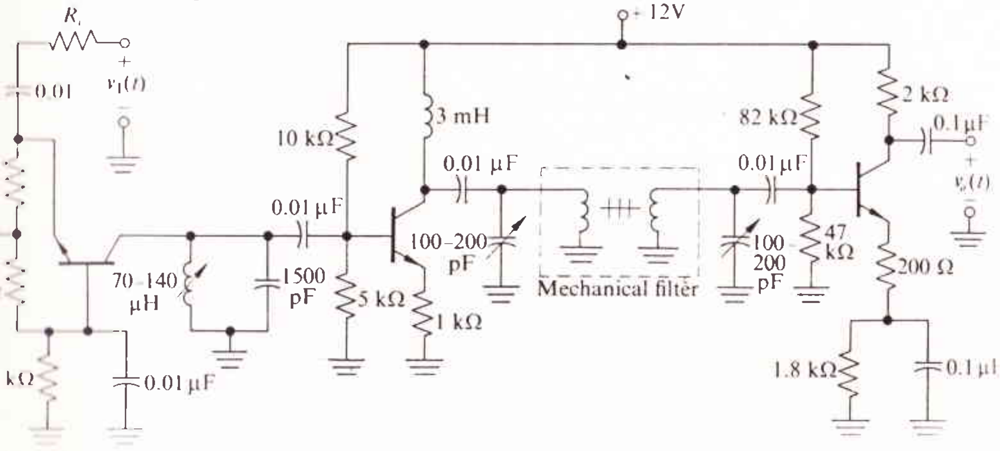
---

† Mechanical filters are mechanically resonant elements whose inputs are driven by, and whose outputs are picked up by, piezoelectric, magnetostrictive, or other electromechanical transducers. The quartz crystal is a mechanical filter; however, as commonly used, the term refers to multi-element (hence multipole) ceramic elements, usually with flat passbands and sharp falloffs in response outside the passband. Chapter 5 of *Solid State Magnetic and Dielectric Devices*, edited by H. W. Katz (John Wiley, New York, 1959), discusses the physical and electrical properties of a number of such devices in some detail.

‡ For those readers with no experience at all with double-tuned circuits, Chapter 18 of E. J. Angelo's *Electronic Circuits* (McGraw-Hill, New York, 1964, second edition) will provide a worthwhile background.



ig. 7.6–1 The amplitude response of a seven-pole commercial mechanical filter in a circuit imilar to that of Fig. 7.6–2.



ig. 7.6–2 Overall IF stage with mechanical filter and preliminary broadband filter to remove purious response possibilities. [Overall gain can be controlled by varying R_i and by capacitively hunting all or part of the 200 Ω in the last emitter. As shown, with R_i = 30 kΩ, the circuit will andle 1 V peak input with an overall gain of 5.]

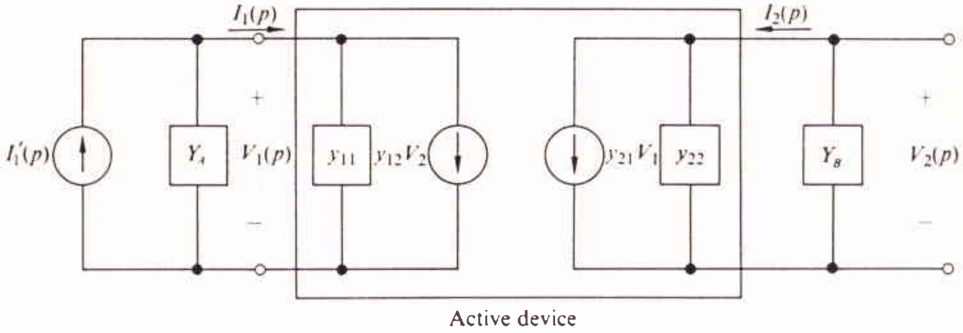


Fig. 7.6-3 Basic y-parameter stage.

are normally measured, since it is desirable that they include all parasitic reactances as well as such difficult-to-model parameters as excess phase shift in transistors.

The y-parameters are usually functions of both Q-point and frequency. Since they are functions of frequency, one must exercise some care when attempting to combine, say, y_{22} from the device and Y_B from the circuit.

In essence, what one must do is make a "small-band" approximation in a normal model. (The "small-signal" approximation is usually already implicit in the y-parameter value.) For example, at frequencies near ω_0 one might approximate y_{22} as

$$y_{22} = g_{22} + j\omega_0 C_{22} + j\delta\omega C'_{22}, \tag{7.6–3}$$

where $\delta\omega = \omega - \omega_0$.

Now if Y_B were a single parallel RLC circuit that combined with y_{22} to tune to ω_0, one would have

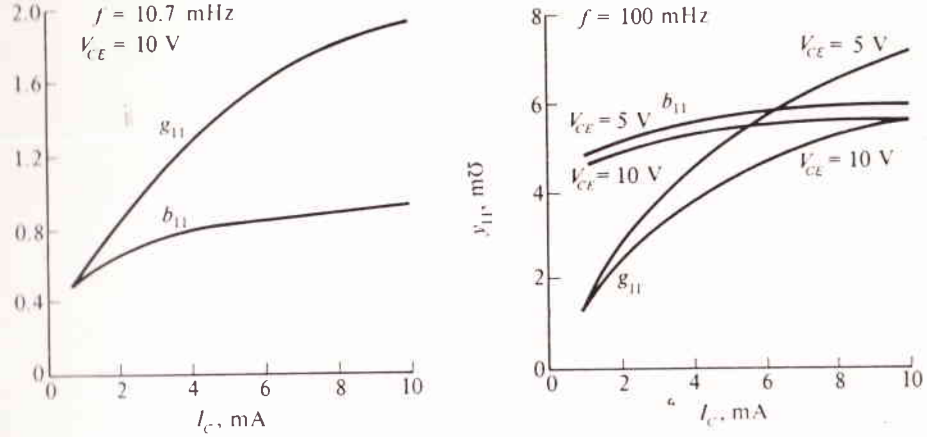
$$Y_t = Y_B + y_{22} = (G_B + g_{22}) + j\omega_0(C_{22} + C_B) - j\frac{1}{(\omega_0 + \delta\omega)L_B} + j\delta\omega(C'_{22} + C_B). \tag{7.6–4}$$

However, if $\delta\omega \ll \omega_0$, then $1/(\omega_0 + \delta\omega) \approx (1/\omega_0) - (\delta\omega/\omega_0^2)$ so that, in the neighborhood of ω_0,

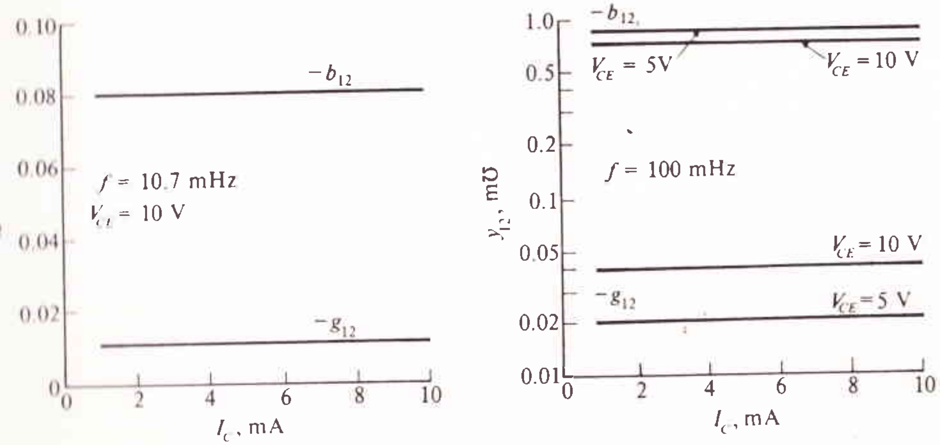
$$Y_t = (G_B + g_{22}) + j\delta\omega\left(C'_{22} + C_B + \frac{1}{\omega_0^2 L_B}\right). \tag{7.6–5}$$

Here, instead of just taking the value of the imaginary part of y_{22} at ω_0 as $j\omega_0 C_{22}$, we have also taken the slope term $j\delta\omega C'_{22}$. In general, there is no reason why C'_{22} should equal C_{22}, just as there is no reason why the incremental slope of the diode equation should equal the ratio of the dc junction voltage to the dc junction current. When the slope of b_{xx} with respect to frequency is small, then $C_{xx} \approx C'_{xx}$. Cases do arise where differences of two or three times occur between the two terms. (One could also take a slope term for the real part of y_{xx}; however, since resonance does not cancel g_{22} or g_{11}, this refinement is seldom necessary.)

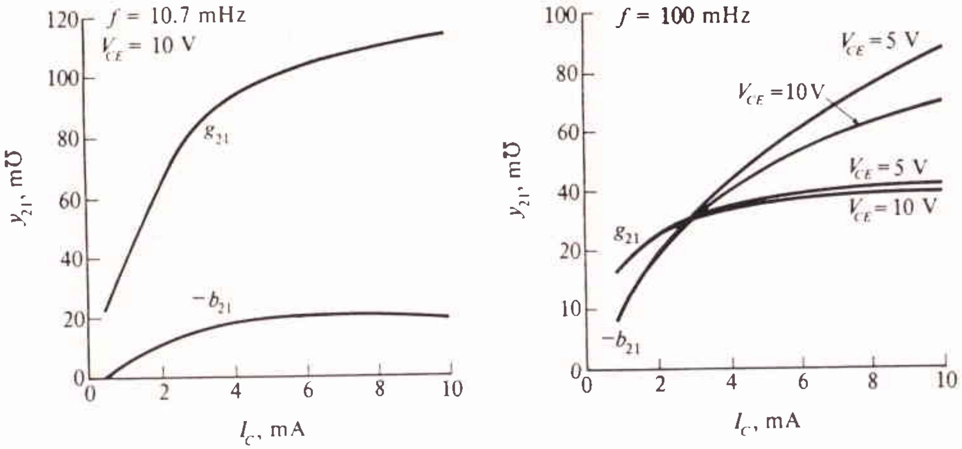
Figure 7.6–4 shows typical 10.7 MHz and 100 MHz common emitter y-parameters vs. collector current for a good NPN silicon bipolar transistor. Figure 7.6–5



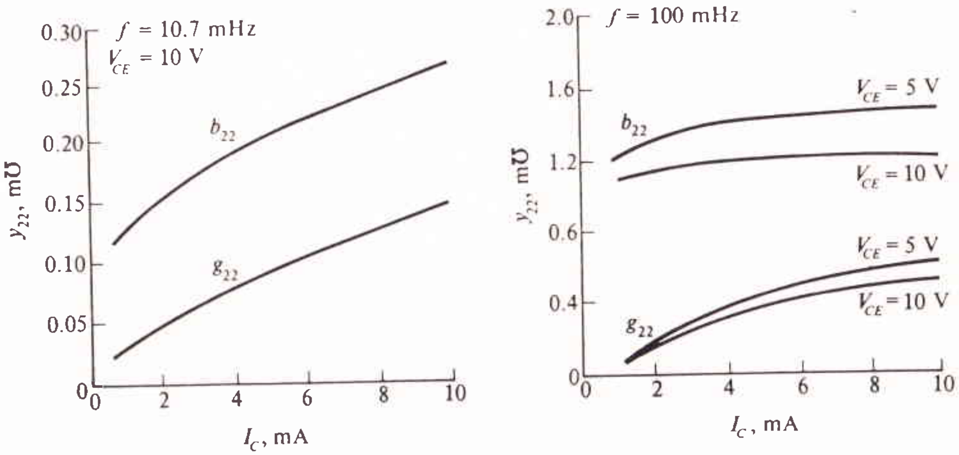
(a) y_{11}: input admittance (output short circuit)



(b) y_{12}: reverse transfer admittance (input short circuit)

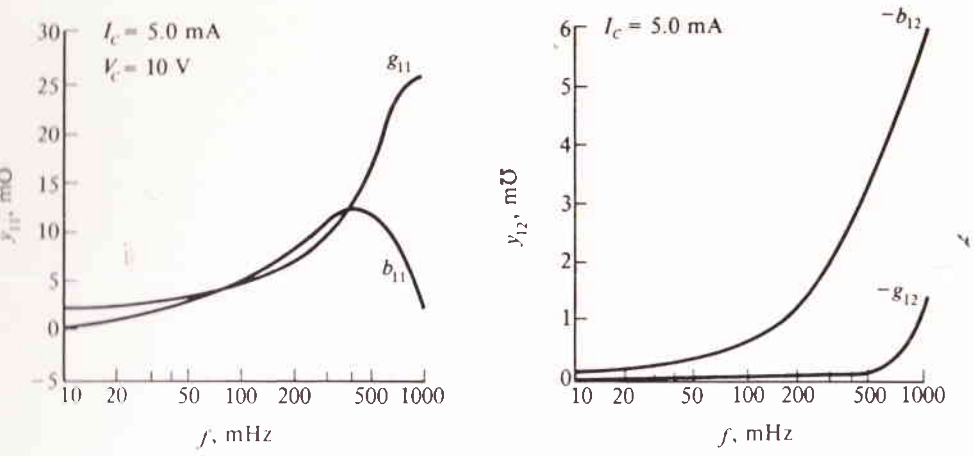


(c) y_{21}: forward transfer admittance (output short circuit)

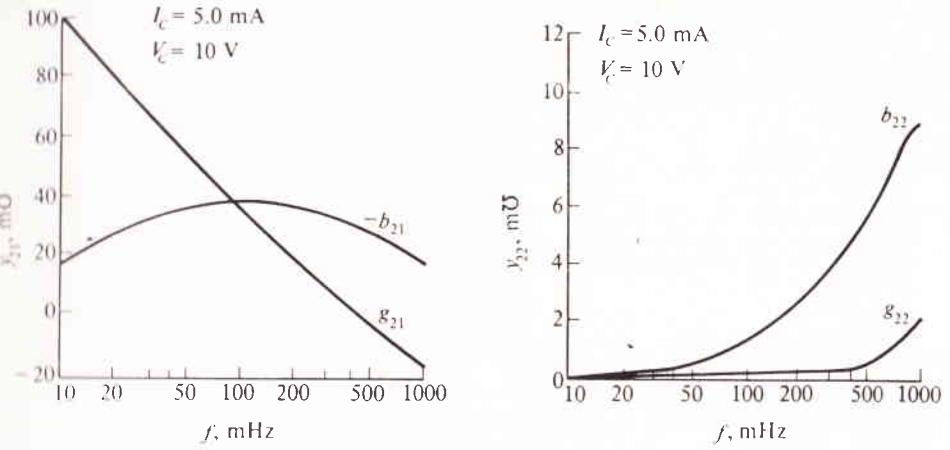


(d) y_{22}: output admittance (input short circuit)

Fig. 7.6-4 Common emitter y-parameters vs. collector current at two different frequencies. 2N918. Reproduced by permission of Fairchild Semiconductor.



(a) y_{11}: input admittance (b) y_{12}: reverse transfer admittance



(c) y_{21}: forward transfer admittance (d) y_{22}: output admittance

ig. 7.6-5 Common emitter y-parameters vs. frequency. 2N918. Reproduced by permission of airchild Semiconductor.

hows the common emitter y-parameters vs. frequency for the same transistor $I_c = 5$ mA, $V_{CE} = 10$ V). It is apparent that these parameters do indeed vary with oth frequency and Q-point, but that one can make a reasonable model over at least ome small range of both quantities.

Consider the problem of finding $Y_{in} = V_1(p)/I_1'(p)$ in the neighborhood of ω_0 hen both the overall input and output circuits are tuned to ω_0 and when both $_A$ and Y_B are simple parallel RLC circuits.

Since $V_2(p) = -y_{21}(p)V_1(p)/[y_{22}(p) + Y_B(p)]$, the effect of the feedback generator $y_{12}V_2$ is to place an impedance in parallel with Y_A and y_{11}. Therefore, we first calculate this "reflected" impedance and then add the y_{11}- and Y_A-terms:

$$Y_{in} = Y_A + y_{11} - \frac{y_{12}y_{21}}{Y_B + y_{22}} = Y_A + y_{11} - y_{12}y_{21}Z_t, \tag{7.6–6}$$

where $Z_t = 1/Y_t$.

If the last term of Y_{in} has a large enough negative real part, then over some range of frequencies, Y_{in} may have a negative real part and the circuit will oscillate in the manner outlined in Section 6.10.

What we wish to demonstrate is a simple graphical approach that not only lets us evaluate $Y_{in}(j\omega_0)$, but also lets us see how $Y_{in}(j\omega)$ varies near ω_0 and exactly what we should do to control this variation.

In general, the practical possibilities in controlling Y_{in} include:

a) **Reduction of** y_{12}. This is always desirable. With a single given FET, vacuum tube, or transistor it can be accomplished by adding an additional out-of-phase feedback to cancel some or all of the device feedback. Such a feedback circuit is known as a neutralizing network.

An alternative is to combine several devices into a circuit such as a cascode circuit or a differential pair and then to consider the y-parameters for the composite circuit. As we shall see, y_{12} for such a circuit may be reduced by a factor of several hundred or more from y_{12} for a single stage.

b) **Increasing** G_B. This corresponds to loading the output circuit. It will reduce the voltage gain of the circuit. If the loading is done by the proper coupling of the impedance from the next stage, then it does not necessarily lead to excessive power loss.

c) **Increasing** G_A. This corresponds to loading the input circuit. Again one is trading gain for circuit stability and symmetry.

Figure 7.6–6 plots $Y_t(j\omega)$ from Eq. (7.6–5) vs. ω in the neighborhood of ω_0. Figure 7.6–7 plots $Z_t(j\omega) = 1/Y_t(j\omega)$.

The straight line in Fig. 7.6–6 is transformed into a circle when one takes its reciprocal to find $Z_t = 1/Y_t$. It will be apparent that the larger G_B is made, the smaller the circle of Fig. 7.6–7 will become. In terms of a magnitude and an angle,

$$Z_t[j(\omega_0 t\,\delta\omega)] = \frac{e^{-j\theta(\delta\omega)}}{G_B + g_{22}},$$

where

$$\theta(\delta\omega) = \tan^{-1}\frac{\delta\omega}{\delta\omega_{3\,\mathrm{dB}}}.$$

The −3 dB passband of the amplifier occupies one-half of the circle; all other frequencies from minus infinity to plus infinity occupy the other half. Actually when

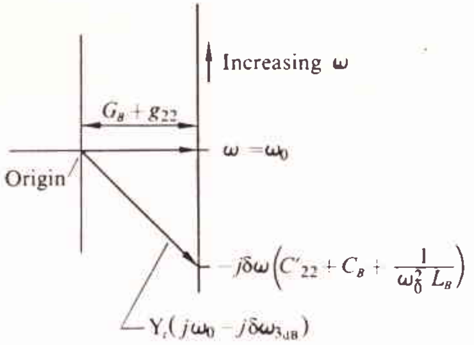
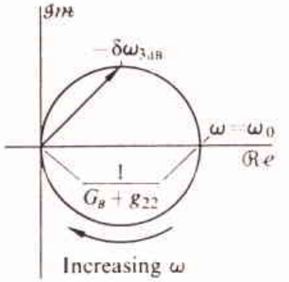

Increasing ω
$G_B + g_{22}$
$\omega = \omega_0$
Origin
$-j\delta\omega\left(C'_{22} + C_B + \frac{1}{\omega_0^2 L_B}\right)$
$Y_t(j\omega_0 - j\delta\omega_{3dB})$

Fig. 7.6–6 Combined circuit and device admittance out at the output of Fig. 7.6–3. Parallel resonance at ω_0 is assumed. Y_t is shown in the neighborhood of resonance.



$$Z_t = \frac{1}{Y_t}$$

Fig. 7.6–7 Reciprocal of the admittance shown in Fig. 7.6–6.

ω is larger, the approximations break down and y_{22} varies with frequency, so that e locus of Z_t will probably cease to be exactly circular. Since the approximation very good in the region of interest and is both adequate and convenient elsewhere, e assume that it holds for all values of ω.

Since resonance is not involved with either y_{12} or y_{21}, we will approximate oth of them by "constant" values across the frequency band in question. Further-ore we will combine the real and imaginary parts of the curves of the type shown Fig. 7.6–4 and 7.6–5 so as to represent y_{12} and y_{21} as magnitudes and angles. With is representation the result of the multiplication shown in Eq. (7.6–6) will be to tate the circle of Fig. 7.6–7 by the sum of the angles of y_{12} and y_{21} and to change the ameter of the circle to $|y_{12}y_{21}|/(G_B + g_{22})$.

For example, if $y_{21} = g_m$ and $y_{12} = -j\omega_0 C_r$, then

$$-\frac{y_{12}y_{21}}{Y_B + y_{21}} = \frac{\omega_0 C_r g_m}{G_B + g_{22}} e^{j[\pi/2 - \theta(\delta\omega)]}, \tag{7.6–7}$$

which is circle 1 in Fig. 7.6–8. Now adding $G_A + g_{11}$ to the reflected admittance will shift the circle to the right by this amount, in addition, if the imaginary part of $Y_A + y_{11}$ is equal to

$$\frac{-j\omega_0 C_r g_m}{G_B + g_{22}},$$

then the circle will be shifted down into position 3 in Fig. 7.6–8.

Curve 3 is no longer quite a circle, since its equation is given by

$$\begin{aligned}\text{Curve } 3 = G_A + g_{11} &+ \frac{\omega_0 C_r g_m}{G_B + g_{22}} \sin\theta(\delta\omega) \\ &+ j\left[\frac{\omega_0 C_r g_m}{G_B + g_{22}} \cos\theta(\delta\omega) - \omega_0(C_{11} + C_A) + \frac{1}{\omega_0 L_A}\right] \\ &+ j\delta\omega\left(C_A + C'_{11} + \frac{1}{\omega_0^2 L_A}\right).\end{aligned} \tag{7.6–8}$$

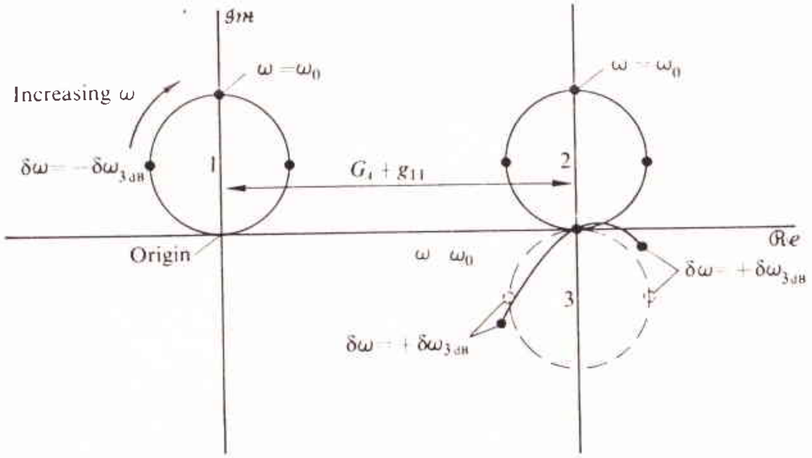

𝔍𝔪
$\omega = \omega_0$
Increasing ω
$\delta\omega = -\delta\omega_{3dB}$
1
$G_A + g_{11}$
2
$\omega = \omega_0$
Origin
$\omega = \omega_0$
ℜe
$\delta\omega = +\delta\omega_{3dB}$
3
$\delta\omega = +\delta\omega_{3dB}$

Fig. 7.6–8 Input admittance of Fig. 7.6–3 for specific y-parameter and circuit conditions, in particular for $y_{21} = g_m$, $y_{12} = -j\omega_0 C_r$, Y_B adjusted to resonate with y_{22}, and Y_A adjusted to resonate with y_{11} and the reflected impedance.

The magnitude and phase of the overall input impedance are given by the magnitude and angle of the vector from the origin to a particular point on curve 3. (Usually circle 3 is a sufficiently good approximation for first-order design purposes.)

It is apparent that, to avoid oscillations, curve 3 must lie completely within the right half of the admittance plane. However, in order to have a stage with a reasonably symmetrical input impedance and with an input impedance that does not become badly mistuned whenever an adjustment is made in the output, it is necessary for the dius of circle 1 to be small in comparison to the distance that circle 2 is to the right of imaginary axis.

is ratio of the radius of circle 1 to $G_A + g_{11}$ is called the alignability factor, k:

$$k = \frac{|y_{12}y_{21}|}{2(G_B + g_{22})(G_A + g_{11})}. \tag{7.6–9}$$

s generally accepted that for well-designed circuits $k \leq 0.20$.

Actually k alone is not really a sufficient specification; for circuits with equal ues for k, those in which circle 1 lies almost completely within the left-half plane l be worse by a factor of $\frac{5}{3}$ than those circuits in which circle 1 lies almost completely the right-half plane.

To repeat our earlier conclusions, one can reduce k either by reducing $|y_{12}|$ or increasing G_B or G_A. (Reducing $|y_{21}|$ is not reasonable, since it reduces the n directly.) One reduces $|y_{12}|$ by switching devices or Q-points, or by a neutralizing lback, or by combining devices into a composite device. Whatever method is d, one must check for the worst-case conditions, because neutralization can mally be achieved only to within a certain tolerance, while the y-parameters y with temperature and as the Q-point is shifted by AVC (or AGC) control signals.

In order to neutralize a stage with a purely capacitive internal feedback, one uires a purely capacitive feedback of opposite phase, as illustrated in Fig. 7.6–9. Normally C_r is not known exactly and varies from unit to unit, with temperature, with Q-point. In addition, one seldom uses better than a 15% tolerance capacitor C_N. Thus it may be reasonable to reduce $|y_{12}|$ to perhaps $\frac{1}{4}$ of its original value, it is unlikely that in a production run one will attempt to do better than this. high-power amplifiers one often does attempt to make adjustments on each lifier so that $|y_{12}|$ is reduced below this value.)

For the purpose of numerical comparison we shall consider an amplifier conng of a single bipolar transistor of the type illustrated in Figs. 7.6–4 and 7.6–5, ell as a cascode and a differential-pair connection of a commercial integrated such as the RCA CA 3005. In each case we arbitrarily set $G_A = g_{11}$. The mon properties are listed in Table 7.6–1. Table 7.6–2 lists the y-parameters and ous derived quantities for all three cases.

Table 7.6–1 Common properties of three amplifiers

$I_{dc} = 2.5$ mA, $|V_{CC}| = |V_{EE}| = 6$ V

Single-tuned, directly connected output circuit

Output coil with $L_B = 8\ \mu$H, $Q = 100$

Output circuit loaded to obtain ±400 kHz bandwidth at $\omega_0 = 6.72 \times 10^7$ Hz; that is, $Q_{\text{loaded}} = 13.4$

Total shunt load from coil losses and device output loss is 7.2 kΩ; that is, $G_B + g_{22} = 139\ \mu$mho for each case

The tuning capacitor for each case is $C_A = (28\text{ pF}) - C_{22}$

Table 7.6–2 List of y-parameters and various derived quantities for three amplifiers (10.7 MHz values; all admittances in μmho; 2.5 mA in single or cascode stages; 1.25 mA in each half of differential pairs)

	Single unit	Cascode†	Differential pair†
y_{11}	$950 + j700$	$2000 + j3000$	$600 + j800$
C_{11}	10.4 pF	48 pF	12.8 pF
C'_{11}	10.4 pF	45 pF	19 pF
y_{12}	$-10 - j80$	$0.22 - j0.08$	$2 - j2$
y_{21}	$75{,}000 - j15{,}000$	$68{,}000 - j12{,}000$	$-20{,}000 + j500$
y_{22}	$60 + j170$	$10 + j150$	$10 + j150$
C_{22}	2.5 pF	2.4 pF	2.4 pF
C'_{22}	2.5 pF	4.8 pF	4.8 pF
y_{12}	$81 \angle -97^\circ$	$0.24 \angle -21^\circ$	$2.8 \angle -45^\circ$
y_{21}	$76{,}000 \angle -11^\circ$	$69{,}000 \angle -10^\circ$	$20{,}000 \angle +178.5^\circ$
$-y_{12}y_{21}$	$6.2 \times 10^{-6}\ \mu\text{mho}^2$ $\angle +72^\circ$	$0.016 \times 10^{-6}\ \mu\text{mho}^2$ $\angle +149^\circ$	$0.056 \times 10^{-6}\ \mu\text{mho}^2$ $\angle -46.5^\circ$
C_B	29.5 pF	29.6 pF	29.6 pF
G_B	101	151	151
$-y_{12}y_{21}/(G_B + g_{22})$	$38{,}500 \angle +72^\circ$	$100 \angle 149^\circ$	$348 \angle -46.5^\circ$
$\dfrac{\lvert y_{12}\rvert \lvert y_{21}\rvert}{2[G_B + g_{22}][2g_{11}]}$	10.1	0.0125	0.146

† The y-parameters of the CA 3005 for the stated conditions are given in Figs. 125 and 126, pp. 176 and 177, of the booklet "RCA Linear Integrated Circuit Fundamentals," Technical Series IC-40, RCA, Harrison, N.J. (1966).

The last entry in Table 7.6–2 is the alignability factor k for the case where $G_A = g_{11}$. Note that when $G_A = g_{11}$ the single-ended case is unstable, while both the other stages have $k \leq 0.146$. The actual difference between the cascode and differential-pair circuits is not as great as the value of k would indicate, since circle 1 in the cascode case is almost completely in the left-half plane while in the differential-pair case it is almost entirely in the right-half plane. To be specific, when $G_A = g_{11}$ and both inputs are tuned to be resistive at ω_0, the net input conductance G_{in} is 3814 μmho in the cascode case and 1448 μmho in the differential-pair case, the difference being only 2.65/1.

To aid in seeing the difference between the two cases as well as in understanding the method, it is desirable that the reader trace out curve 3 for both cases. From this curve one can get an idea of the input impedance symmetry for each circuit.

It is sometimes useful to compare amplifying stages in terms of their power gains, since power does not have the indefiniteness that is sometimes associated with voltage or current gains.

A straightforward analysis of Fig. 7.6–3 will show that the resonant power gain defined as the load power divided by the power supplied by the I'_1 generator is

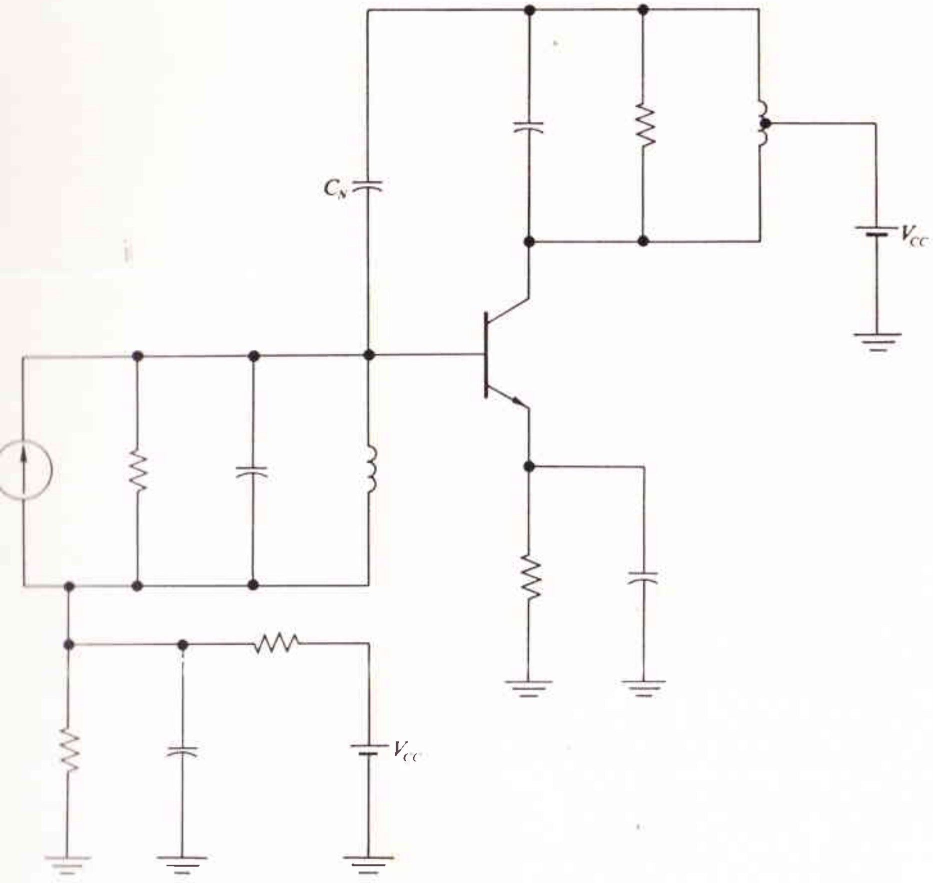


Fig. 7.6–9 Possible neutralization circuitry when $y_{12} = -j\omega C_r$.

en by Eq. (7.6–10):

$$A_P = \frac{P_{\text{load}}}{P_{\text{gen}}} = \frac{|y_{21}|^2 G_B}{(G_A + g_{11} + g_R)(g_{22} + G_B)^2}. \tag{7.6–10}$$

This equation presumes that the input and output are both tuned to resonance; is the resistive portion of the impedance reflected from the output through the generator, and G_B is assumed to be the desired load resistor.

For the purpose of comparing stages, we consider a special form of A_P in ich $g_R = 0$ and both the input and output are "matched" resistively. For this case, becomes $A_{P_{\text{mat}}}$ and Eq. (7.6–10) reduces to

$$A_{P_{\text{mat}}} = \frac{|y_{21}|^2}{8 g_{11} g_{22}}. \tag{7.6–11}$$

Now we compare $A_{P_{\text{mat}}}$ for the three stages and A_P for the last two cases, suming $g_{11} = G_A$, but allowing g_R, G_B, and g_{22} to have their previous values.

(Since the single-ended transistor without neutralization will oscillate, a calculation for A_P is meaningless.)

	Single unit	Cascode	Differential pair
$A_{P_{\text{mat}}}$	41.1 dB	44.5 dB	39.2 dB
$A_{P_{\text{actual}}}$	—	38.5 dB	35.3 dB

From these results we see that, while neither of these composite stages yields its "matched" power gain, two stages of either composite type would yield a theoretical gain more than 70 dB and a practical gain of at least 60 dB.

We might attempt to salvage the single-ended stage by neutralizing it. If we assume that we can neutralize the reactive portion of y_{12} to within 10%, then

$$\mathscr{R}e[y_{12}] = -10, \qquad +8 > \mathscr{I}m[y_{12}] > -8,$$

from which, for the two extreme cases, we obtain the following values:

	$y_{12} = -10 + j8$	$y_{12} = -10 - j8$
y_{12}	$12.8 \angle +141.5°$	$12.8 \angle -141.5°$
$y_{12}y_{21}$	$0.98 \times 10^{-6}\ \mu\text{mho}^2 \angle -49.5°$	$0.98 \times 10^{-6}\ \mu\text{mho}^2 \angle +27.5°$
$-y_{12}y_{21}/(G_B + g_{22})$	$6070 \angle -49.5°$	$6070 \angle +27.5°$

If one constructs the type 1 circles indicated by the values above, one will see that the $y_{12} = -10 + j8$ case is the worst case but that, even here, adding 1100 μmho of real input conductance (via g_{11} and G_A) will yield absolute stability; hence, if $g_{11} = G_A = 950\ \mu$mho, then the type 3 circles will lie completely within the right-half plane. The alignability values will not be less than 0.20. However, since the circles are not rotated so far from the real axis to begin with, the variation across the band may not be excessive. If this variation still seems too much, the next step would probably be to reduce the load seen at the output. This can be done without changing the bandwidth by using a step-up transformer between the transistor output and the LC circuit. For example, with an effective turns ratio of 1:2 we may increase G_B to 634 μmho; this reduces the diameter of circle 1 to 1517 μmho and makes $k = 0.40$. The result, with the small angle of rotation of circle 1, should be a usable stage.

Note that this deliberate mismatching at the output has decreased the power gain while increasing the alignability. This trade is often a reasonable one to make. For this neutralized, heavily loaded stage, as specified for the $y_{12} = -10 + j8\ \mu$mho case, the actual resonant power gain is 3100 or 34.9 dB; hence the single-ended case is still competitive with the cascode and differential-pair cases if gain alone is considered. Thus by partial neutralization and the use of sufficient input or output mismatching we can obtain reasonably alignable single-unit stages with power gains of the order of 30 dB per stage.

.7 STAGES WITH DOUBLE-TUNED CIRCUITS

igure 7.7–1 illustrates three possible types of coupling between small-signal double-ined circuits. Obviously a great many variations are possible with such circuits. o illustrate the basic concepts we shall restrict ourselves to the specific cases out-ned below:

) each half-circuit (uncoupled) identical and tuned to ω_0;
) only a single type of coupling in a given circuit;
) overall bandwidths of 5% or less of ω_0.

or all such cases where $Y_1 = Y_2$†,

$$\frac{V_2(p)}{I_1(p)} = \frac{Y_m(p)}{Y_1(p)[2Y_m(p) + Y_1(p)]},$$

$$\frac{V_1(p)}{I_1(p)} = \frac{Y_1(p) + Y_m(p)}{Y_1(p)[2Y_m(p) + Y_1(p)]}. \tag{7.7–1}$$

The transfer impedance for such cases always has two sets of complex poles.)ne set is at the zeros of $Y_1(p)$ and hence has a real part of $-1/2R_1C_1$ and a resonant equency of $1/\sqrt{L_1C_1}$. The quadratic for the other set of poles is the one resulting om the $[Y_1(p) + 2Y_m(p)]$-term; hence it will vary with the type of coupling. Table 7–1 lists the results for the three cases under consideration.

able 7.7–1 Comparison of three types of coupling

	Inductive	Mutual inductive	Capacitive
uadratic	$p^2 + \frac{p}{R_1C_1} + \frac{1+2k}{LC}$	$p^2 + \frac{p}{R_1C_1} + \frac{1}{(1+k)LC}$	$p^2 + \frac{p}{R_1C_1(1+2k)}$ $+ \frac{1}{LC(1+2k)}$
cale factor	L_1/L_m $k\omega_0^2 C_1$	M/L_1 $\omega_0^4 k L_1(1-k^2)$	C_m/C_1 $k(1+2k)C_1$
eros at origin	1	1	3

Actually for mutual inductive coupling the convenient formulation puts the lenominator in a form in which both sets of poles are functions of k. The normal pproximation for small values of k is to ignore the variations in Q caused by k (that s, to neglect the horizontal motion of the pole in the capacitive case) and to consider he vertical movement. For $k \ll 1$,

$$\frac{1}{\sqrt{1+k}} \approx 1 - \frac{k}{2} \quad \text{and} \quad \frac{1}{\sqrt{1-2k}} \approx 1 - k.$$

E. J. Angelo (*Electronic Circuits*, McGraw-Hill, New York, 1964, second edition) illustrates the raightforward extension to the case where $Y_2(p) = aY_1(p)$.

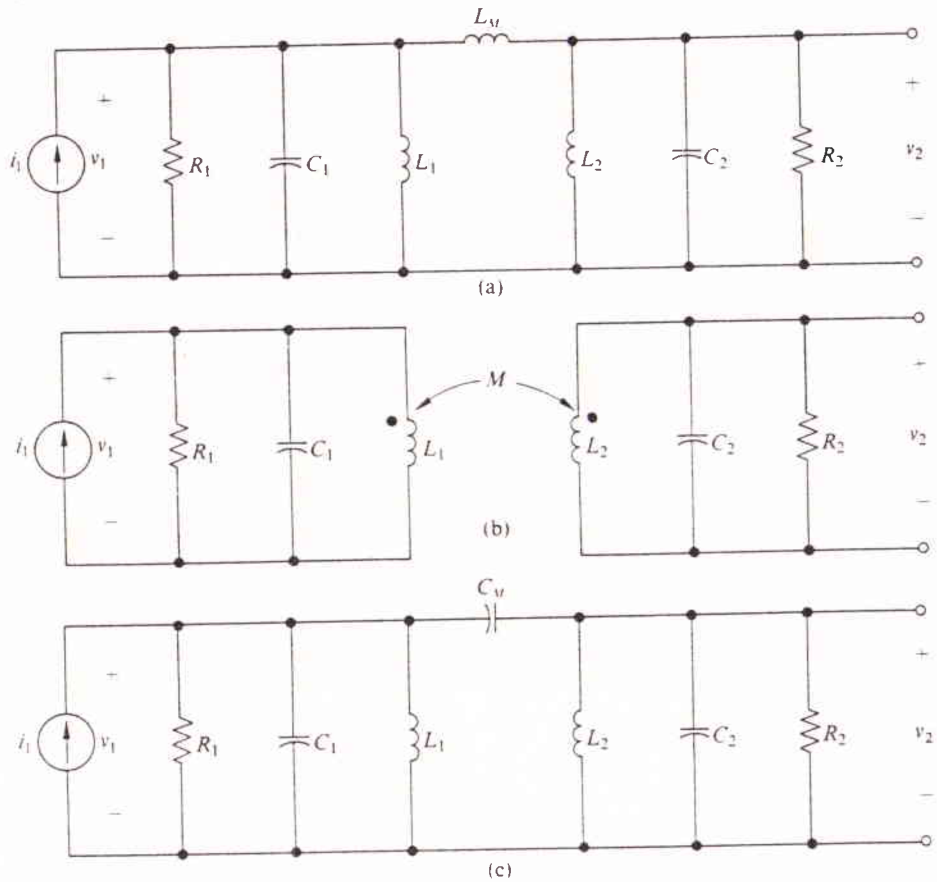


Fig. 7.7–1 (a) Inductive coupling. (b) Mutual inductive coupling. (c) Capacitive coupling.

The pole motion for the upper set of complex poles for the three cases is shown in Fig. 7.7–2. It is apparent that, except for the skewness caused by the variations of Q with k, all three cases may be used to obtain maximally flat or other two-pole network adjustments. Though the inductive case has the easiest algebra, the other two cases turn out to be much more practical. Both of these cases are widely used. For maximal flatness in the capacitive case, $\omega_0 C_m R = 1$; in the mutual inductive case, $\omega_0 MR = 1$.

The question that naturally arises is: How does the input impedance of a double-tuned circuit affect the stability and alignability calculations of the previous section? From Fig. 7.6–6 it is clear that the admittance reflected into the input side of the y-parameter model is always $-y_{12}y_{21}Z_t$, where Z_t is the total impedance across the y_{21} generator. That is,

$$Z_t = \frac{1}{y_{22} + Y_B}.$$

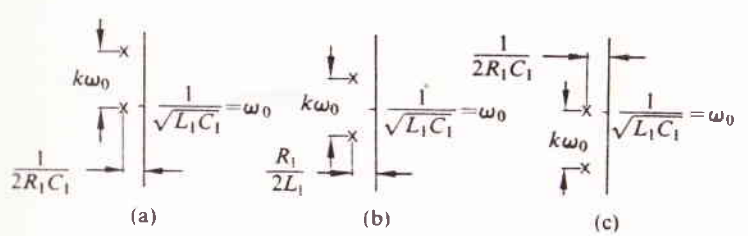


Fig. 7.7–2 (a) Inductive coupling. (b) Mutual inductive coupling. (c) Capacitive coupling.

If we assume that y_{22} is absorbed into the primary tuned circuit, then Z_t is the input impedance for the double-tuned circuit. Equations (7.7–2) and (7.7–3) list Z_t for the mutual inductive and capacitive coupling cases:

Mutual inductive coupling

$$Z_t = \frac{p\left(p^2 + \dfrac{p}{R_1C_1} + \dfrac{\omega_0^2}{1-k^2}\right)}{C_1\left(p^2 + \dfrac{p}{R_1C_1} + \dfrac{\omega_0^2}{1+k}\right)\left(p^2 + \dfrac{p}{R_1C_1} + \dfrac{\omega_0^2}{1-k}\right)} \tag{7.7–2}$$

Capacitive coupling

$$Z_t = \left(\frac{1+k}{1+2k}\right)\frac{p\left[p^2 + \dfrac{p}{R_1C_1(1+k)} + \dfrac{\omega_0^2}{1+k}\right]}{C_1\left(p^2 + \dfrac{p}{R_1C_1} + \omega_0^2\right)\left[p^2 + \dfrac{p}{R_1C_1(1+2k)} + \dfrac{\omega_0^2}{1+2k}\right]}. \tag{7.7–3}$$

Though these equations look rather formidable, both can be greatly simplified if we consider only the region near $\omega_0 = 1/\sqrt{L_1C_1}$ and if we maintain our assumption of $k \ll 1$. In this case both equations have two poles that are approximately equally spaced about the zero. If we consider the maximally flat transfer function case, then Z_t at the center frequency [ω_0 in the mutual inductive case and $\omega_0(1 - k/2)$ in the capacitive coupling case] is $R/2$, while the approximated normalized values of the impedance in terms of $\alpha = 1/2R_1C_1 = \omega_0/2Q_T$ are given by Table 7.7–2. The center-frequency transfer impedance for the same case is $R/2(1 - k^2)$, which is approximately $R/2$ and thus approximately the same as Z_t.

Figure 7.7–3 plots the normalized data of Table 7.7–2 and compares them with a circle of radius $\frac{1}{2}$. That is, Fig. 7.7–3 compares the data of Table 7.7–2 with the input impedance to be expected from a single-tuned circuit having the same center frequency and a load R. However, this is just the case of the single-tuned circuit.

The use of the same graphical approach to calculate the complete reflected impedance and the resultant alignability is a perfectly straightforward procedure. However, it is apparent that if the single-tuned case with load R is stable, then, because the new impedance lies within the circle of the single-tuned case, the double-tuned, maximally flat case with a load of R per side will also be stable.

Table 7.7–2 Normalized input impedance for identically tuned coupled circuits (R_1 on each side)

$\delta\omega/\alpha$	$\mathscr{R}e(Z_t/R)$	$\mathscr{I}m(Z_t/R)$	$\lvert Z_t\rvert R$	Angle
0	0.50	0.00	0.50	0°
+1	0.60	∓0.20	0.635	∓18.5°
±2	0.259	∓0.40	0.50	∓53°
±4	0.070	∓0.25	0.26	∓74°

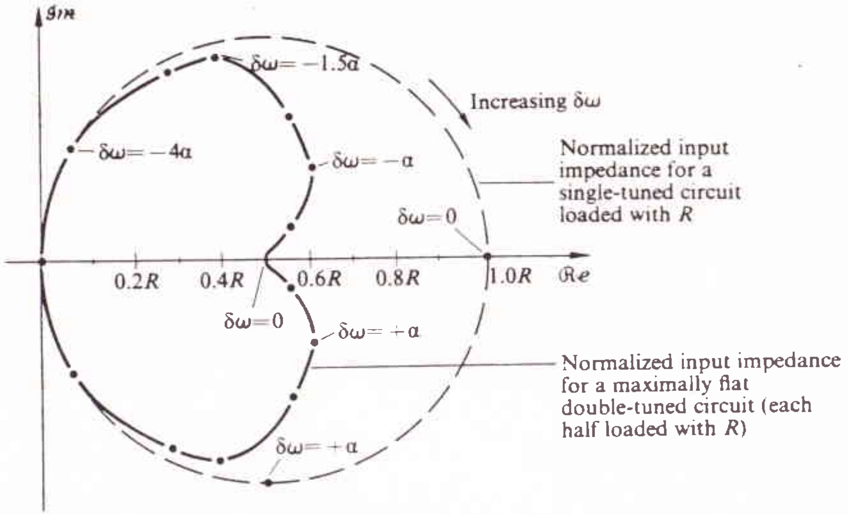


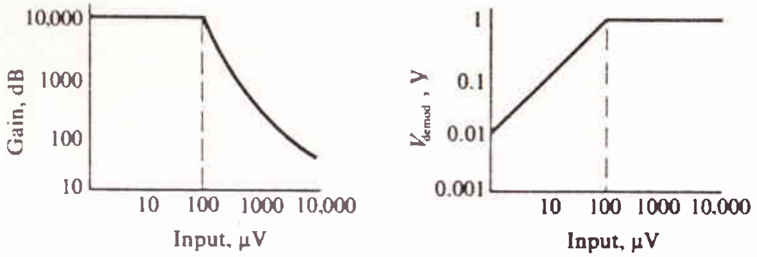
Fig. 7.7–3 Plot of the normalized Z_{in} for a maximally flat double-tuned identical set of tuned circuits.

7.8 GAIN CONTROL CIRCUITS

Since the level of the signals present at the input of a receiver can vary by more than $10^6/1$ while the "linear" range of operation of most amplifiers is $10^3/1$ or less and the practically useful range of many demodulation circuits is only about 10/1, it is imperative that most receivers have a built-in automatic gain control circuit.

Such a circuit should sense the long-term average (long-term with respect to the lowest modulation frequency) value of the signal and adjust the amplifier gain in such a manner that all stages operate within their optimum signal level ranges. Ideally, as the input signal level increases, the gain should remain at its maximum value until the optimum demodulator input level is reached, and then for further increases in the input the gain should decrease in such a fashion that the output carrier level remains constant.

Figure 7.8–1 illustrates the ideal situation for the case where the optimum demodulator voltage input is 1 V and the maximum amplifier gain is 80 dB. Practical



(a) (b)

Fig. 7.8–1 (a) Ideal gain vs. input signal for AGC amplifier. (b) Ideal $v_o(t)$ vs. $v_i(t)$ for AGC amplifier.

amplifiers usually depart from the ideal situation both at low input levels, where the gain begins to decrease sooner than it should, and at high levels, where some amplifiers become seriously overloaded.

Another practical problem with most gain control schemes is that they modify the active device's input and output impedance and hence cause detuning and/or bandwidth changes. In addition, many of the schemes modify the circuit's noise and distortion properties in what may turn out to be an unsatisfactory manner. Thus the best gain control circuit is usually a compromise among a number of conflicting requirements.

Before we can intelligently discuss overall systems, we need to outline the possible gain control mechanisms. As we shall see, there may be several possible means of varying the gain of a given device. These may vary with the frequency of operation. In some cases several identical-looking circuits may in fact operate in somewhat different manners, since they operate at widely different center frequencies.

The most straightforward method of gain control is to vary g_m in the device model shown in Fig. 7.8–2. Such a simplified model can represent a vacuum tube, a junction FET, a MOS FET, or a bipolar transistor in the frequency range where both the base spreading resistor r_{BB} and the base recombination resistor $r_{B'C}$ or r_π may be neglected as we have been doing in previous sections.

Vacuum tube pentodes exist in which g_m varies smoothly (but not linearly) from, say, 5000 µmho at zero grid-to-cathode voltage to zero micromhos at a V_{GK} value of -30 V. In other tubes the same g_m range may be covered in only a 5 V swing of the bias voltage.

The wide-range or "remote-cutoff" device is made by effectively paralleling (in the same envelope) several tube structures. One structure has a high g_m at zero voltage but cuts off at a relatively low grid-to-cathode voltage, while another has a low g_m at zero voltage and decreases only slowly with increasing bias. The advantage of such a control characteristic is that, at large biases and low g_m's and hence low gains, the g_m is nearly constant over a wide voltage range; hence the device will be able to handle large signals without distortion.

Figure 7.8–3 illustrates the g_m vs. V_{GK} for a pentode vacuum tube, as well as the gain for a two-stage tandem of such pentodes in which $R_L = 10\,\mathrm{k}\Omega$ is assumed for each stage. The amplifier illustrated does have a range of controlled gain of more than $10^6/1$. Whether this whole range is usable is a question we must now examine.

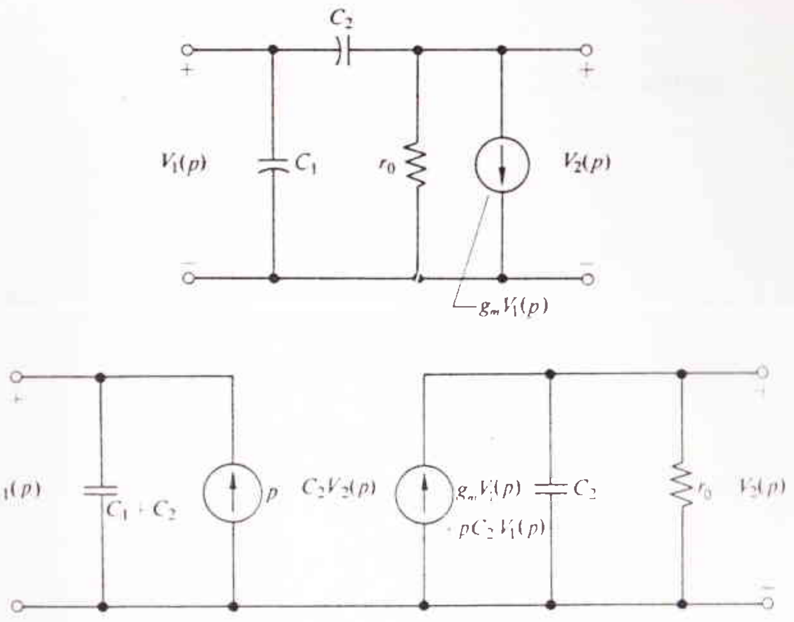


Fig. 7.8–2 (a) π-model for a device. (b) Equivalent y-model for the device of Fig. 7.8–2(a).

Even if g_m is set to zero in the models of Fig. 7.8–2, the remaining voltage "gain" of a single stage has a magnitude of $1/\sqrt{1 + [(g_0 + G_L)/\omega C_2]^2}$ where G_L is the load conductance. For large values of ω this term approaches unity. If $C_2 = 1$ pF, $\omega = 60 \times 10^6$, and $g_0 + G_L = 100$ µmhos, then the resultant "gain" is 0.51 rather than zero. If two stages are cascaded, then the overall gain depends not only on how the g_m of each stage is controlled, but also on the composite model. The overall result is that the capacitive feedthrough term prevents one from reaching the lower portions of the characteristic of Fig. 7.8–3 at high frequencies. This capacitive effect is reduced in the same way that it was reduced in Section 7.6, that is, by neutralization.

In almost all AGC circuits, proper neutralization will extend the lower-gain end of the curve by the order of 20 dB. The limit to such extensions is the accuracy of the neutralization of a particular stage. In addition to the normal tolerance problems, the AGC action is likely to vary the internal feedback parameters, and at low gains the large input signals present may effectively vary the "small-signal" model parameters.

In both junction and MOS FET's, g_m is proportional to $\sqrt{I_D}$; hence one can obtain g_m vs. $\sqrt{I_D}$ (or g_m vs. $V_{GS} - V_{Th}$ curves for MOS transistors or g_m vs. $V_p - V_{GS}$ curves for junction FET's). In either case it is possible for the manufacturer to build "remote-cutoff" devices or for the user to parallel several units to obtain this effect. In theory, one can obtain results similar to those for vacuum tubes, although probably

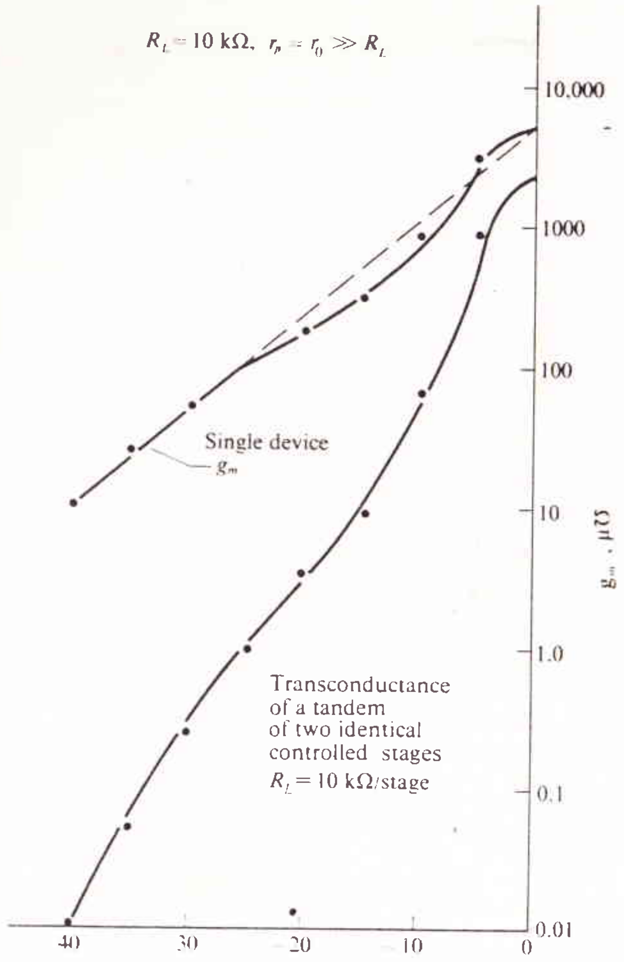


Fig. 7.8–3 g_m and two-stage gain vs. V_{GK} (vacuum-tube gain-controlled amplifier).

ith a somewhat reduced gate-source range. Again neutralization will be necessary reach the lower gain levels.

For bipolar transistors operated in the frequency range where Fig. 7.8–2 is a valid epresentation ($\omega C_1 r_{B'C} \gg 1$) *and* with a low enough bias current that the emitter-ase transition region capacitance dominates C_1 ($C_1 = C_\pi = C_{B'C}$), g_m is proportional I_E and thus there is a linear relationship between gain and emitter current. For any high-frequency transistors with small effective diffusion capacitances, bias urrents below 1 mA lead to the desired operation.†

---

or higher currents C_1 becomes a linearly increasing function of I_E; hence the current gain ecomes a "constant" as the decrease in the input impedance tends to compensate for the increase g_m. In addition, the variation of C_1 with I_E may cause serious detuning problems.

A number of other types of gain control circuits are possible and are in use.

All transistors have alphas and betas that are to some extent functions of the dc bias current. Normally the transistor designer tries to minimize these changes; however, in an AGC situation he may try to utilize them.

One approach is represented by the "tetrode" transistors, in which a second base connection is added to the device. The second base current now controls the alpha or beta of the "normal" transistor. Beta variations of 20/1 or more should be possible for a 10/1 variation of control current.

A second approach uses the fact that the gain of certain bipolar transistors falls off, not only as I_E increases, but also as V_{CB} decreases, particularly when the absolute value of V_{CB} is small. In such transistors (a 2N2415 at 70 MHz, for example), by varying V_{CB} from 1.2 V to 2.5 V (this is equivalent to a 1.3 mA change across a 1 kΩ resistor) one can change the gain by a factor of 100 (40 dB). For larger values of V_{CB} the gain will be reasonably independent of V_{CB} (or V_{CE}). A circuit using this phenomenon for gain control purposes is known as a "forward-acting AGC" circuit.

Since variation of V_{CB}, especially for small values of V_{CB}, causes a large variation in $C_{B'C}$, it will be difficult to achieve neutralization with "forward-acting" AGC of the variable-voltage type.

All bipolar gain control circuits in which I_E is varied actually operate by means of a combination of the mechanisms outlined previously plus mismatching effects at both the input and output of the device. Figure 7.6–4 clearly indicates such input and output admittance variations at 10.7 MHz and 100 MHz. At lower frequencies the variations in output impedance may become less important; however, the variation of input impedance with I_E continues down to dc.

In the usual tuned amplifier, one does not want to depend on mismatching for gain control, since mismatching also causes detuning and selectivity changes; hence one deliberately taps the input and output transformers down to the point where the circuit not only is stable and alignable but is also relatively immune to shifts in tuning caused by AGC action.

The gains of cascode connections are varied as a common emitter stage, while the g_m of the differential-pair configuration is controlled by varying I_k.

Another approach to the AGC or AVC problem is to use a separate attenuator section or to shunt one of the tuned circuits by a variable attenuator. Possibilities include an FET or light-controlled resistor as illustrated in Figs. 6.9–3 and 6.9–4 or a diode attenuator as illustrated in Figs. 5.5–7, 5.5–9, or 7.4–3.

In any case, one must derive a feedback voltage (or current) to drive the AGC circuit from the signal itself. This is normally done by means of one of the detector circuits of Chapter 10. One should note that the desired relationship between the AGC voltage (or current) and an increase in the signal amplitude depends on the type of AGC employed. It should also be apparent that the desirable feature of a "delay" in the start of gain reduction until a desired minimum output is obtained may be incorporated either in the detector circuit or in the gain control circuit. For example, bipolar circuits that depend on a reduction of V_{CB} do not produce large changes in gain until V_{CB} falls below about 2.5 to 3 V. Hence if V_{CC} is 9 V, if there is a 1.5 kΩ dropping resistor, if $V_B = 0$, and if $I_{EQ} = 3$ mA, then the first 1 mA change in I_E will not cause much change in gain, while the second 1 mA change in I_E will cause a large change in the gain.

It should also be apparent that, since the AGC system is a feedback system, the loop filtering must be properly designed or oscillations might occur. In particular, excessive low-frequency filtering must not be added to remove the modulation signal from the control signal path.

7.9 NOISE, DISTORTION, AND CROSS MODULATION

Noise and cross modulation are problems only in the early stages of a receiver; if these problems are not solved there, no later stage can correct them. Nonlinear distortion may occur in any stage, but once adequate filtering has removed any interfering carriers, this type of distortion is most likely in the last stage of the amplifier where the signal level is a maximum.

The techniques of Chapter 4 indicate, for various devices, what signal levels are possible without such nonlinear distortion and what form it will take when it occurs. We shall not discuss it further here.

Noise has several prime sources. One is the antenna; noise arriving from this source must be removed by adequate filtering. If all the earlier stages operate linearly and if noise terms at all the potential interfering frequencies are adequately removed, then the IF selectivity will remove the out-of-band portions of such noise. The second component of noise is in-band noise generated in the receiver itself. Such noise is produced in the FET, the bipolar transistor, or the vacuum tube. While every device has its own peculiarities, the noise normally has both an input circuit and an output circuit component and it normally increases with the bias current. In general, antenna noise dominates below about 30 MHz and front-end receiver noise dominates at higher frequencies.

One way to compare amplifiers with respect to their noisiness is by the noise figure, F:

$$F = \frac{\text{Signal/Noise Ratio at Input}}{\text{Signal/Noise Ratio at Output}}. \tag{7.9–1}$$

For an ideal noiseless amplifier, F is unity or zero dB, since the amplifier adds no noise of its own. Practical amplifiers always have $F > 1$. (Obviously both input and output signal/noise ratios need to be measured in the same bandwidth; otherwise, values of F apparently less than one can be obtained.)

In addition to the noise produced in the desired band itself, both antenna and receiver noise outside the band may get translated into the band by direct modulation of an amplifier, by cross modulation in an amplifier, or by being at the appropriate mixer interference point. Modulation by low-frequency noise (which is often an order of magnitude larger than in-band noise) is avoided by providing adequate low-frequency bypassing and by operating the initial amplifier in a linear manner.

Since the amplifier noise figure is related to the amplifier's power gain, it will be a function of the source impedance seen by the device. Normally there is a level of source impedance that is optimum from the viewpoint of the noise figure. Fortunately, this minimum in F as the source impedance is varied is fairly broad; hence 2/1 noise mismatching should cause less than a 1 dB increase in F. For a junction or MOS transistor the optimum source impedance is of the order of 2 kΩ, for a bipolar transistor 200 Ω is apt to be much more reasonable.

Since the erratic division of current between the grids makes multigrid tubes much noisier (perhaps five times noisier), one uses triode vacuum tubes as RF amplifiers in frequency regions where receiver noise is important. Of course, a "good" pentode may still be better than a noisy triode, so discrimination must still be exercised here.

A good RF pentode at 60 MHz might have an optimum source impedance of 3 kΩ and a noise figure of 4 dB, while the same tube connected as a triode might have an optimum source impedance of 1.4 kΩ and a noise figure of 2 dB.

Cross modulation can occur whenever the series expansion for the output current as a function of the input voltage contains a cubic term. For example, if the output current of an amplifier has the form

$$i_o(t) = I_{dc} + g_m v(t) + a_2 v^2(t) + a_3 v^3(t), \tag{7.9–2}$$

where $v(t)$ is assumed to have the form

$$v(t) = V_1 \cos \omega_1 t + V_2[1 + m_2 f(t)] \cos \omega_2 t \tag{7.9–3}$$

$[|f(t)| \leq 1$ and the highest radian frequencies in $f(t)$ are much less than ω_1 or $\omega_2]$, then there will be output terms with center frequency ω_1 as shown in Eq. (7.9–4):

$$i_{\omega_1}(t) = \{g_m V_1 + \tfrac{3}{2} V_1 V_2^2 [1 + m_2 f(t)]^2 a_3\} \cos \omega_1 t. \tag{7.9–4}$$

Thus even if the original signal at ω_1 is unmodulated, an interfering signal at ω_2 may transfer its modulation to the desired signal frequency.

If $f(t) = \cos \mu_2 t$, then the transferred modulation will have sideband components at both μ_2 and $2\mu_2$. If m_2 is small enough, say 0.10, one might ignore the second, harmonic term and ask for the size of V_2 that would cause 1% sinusoidal modulation on the desired carrier. From Eq. (7.9–4) this value would be

$$V_2^2 3 m_2 \frac{a_3}{g_m} = 0.01, \tag{7.9–5}$$

which is independent of V_1 and of ω_2. Thus, unless a_3 can be made very small with respect to g_m, adequate selectivity must be provided to keep V_2 at the device input below the desired level. For example, if $|a_3/g_m| = 1$ and $m_2 = 0.10$, then $V_2 \leq 180$ mV.

An ideal FET amplifier, with no series resistance, has $a_3 = 0$ and hence no cross modulation problems. Ideal differential pairs also have $a_3 = 0$ and again should not cause cross modulation. Practical FET's and differential pairs will have nonzero but small values for a_3; hence they should have only very small cross-modulation effects.

The cross-modulation term for the bipolar transistor can be calculated from (7.2–7) if we assign ω_s to the desired signal and ω_0 to the interfering signal; then

$$i_{\omega_s}(t) = 2I_{ES}e^{qV_{dc}/kT}I_0(x)I_1(y)\cos\omega_s t, \qquad (7.9\text{–}6)$$

iere $y = V_1 q/kT$ is the normalized version of the desired signal and

$$x = \frac{qV_2[1 + f(t)]}{kT}$$

the normalized version of the undesired signal. For small values of x and y we n expand $I_0(x)$ as $I_0(x) \approx 1 + (x^2/4)$ and $I_1(y)$ as $I_1(y) \approx y/2$ so that $i_{\omega_s}(t)$ becomes

$$i_{\omega_s}(t) = I_{ES}e^{qV_{dc}}\frac{V_1 q}{kT}\left\{1 + \left(\frac{qV_2}{kT}\right)^2\frac{[1 + f(t)]^2}{4}\right\}\cos\omega_s t. \qquad (7.9\text{–}7)$$

If $f(t) = m_2 \cos\mu_2 t$, then for small values of m_2 the value of V_2 that will cause cross modulation of the desired carrier is approximately

$$\left(\frac{qV_2}{kT}\right)^2\frac{m_2}{2} \leq 0.01. \qquad (7.9\text{–}8)$$

r $m_2 = 0.10$, V_2 must be less than 8.6 mV.

In addition to cross modulation, the cubic term also causes distortion of the dulation originally on the desired carrier. For example, if

$$v_1(t) = V_1[1 + g(t)]\cos\omega_1 t,$$

en the cubic term causes a current component at ω_1 of

$$\tfrac{3}{4}V_1^3\{1 + 3g(t) + 3[g(t)]^2 + [g(t)]^3\}\cos\omega_1 t. \qquad (7.9\text{–}9)$$

viously such a term will cause distortion in a single sine-wave modulation and ermodulation in any multifrequency modulation.

As a practical matter, 1% cross modulation is barely visible in a television picture d easily tolerable in voice communication. With 5% cross modulation a television ture is badly distorted, although voice communication is still possible without lue difficulty.

OBLEMS

Assume a mixer of the type shown in Fig. 7.2–1(a) with a silicon FET having $I_{DDS} = 4$ mA, $V_p = -4$ V, and an ac bypassed source resistor of 2 kΩ. The oscillator peak voltage is 1.8 V. Evaluate the conversion and amplifier transconductance from the gate-ground terminals to the drain current. For an input sinusoidal signal of 1 mV peak amplitude, determine the drain current components at the input signal frequency, at the oscillator frequency, and at the difference frequency.

In the mixer of Problem 7.1, what is the limit on the size of an interfering signal input before nonlinearities result in the transfer of amplitude modulation from this signal to the difference frequencies? Explain.

7.3 Consider a bipolar mixer of the type shown in Fig. 7.2–3. Assuming that the dc component of the emitter current is 1 mA and that the oscillator drive voltage (across the emitter side of the driving transformer) is 100 mV peak, determine G_c and G_m. Assuming that the input signal has a peak amplitude of 2 mV, evaluate the collector current component at the signal, the oscillator, and the difference frequencies.

7.4 Estimate the effect on the results for the mixer of Problem 7.3 of adding an unbypassed 20 Ω resistor in series with the transistor's emitter while leaving all other quantities unchanged.

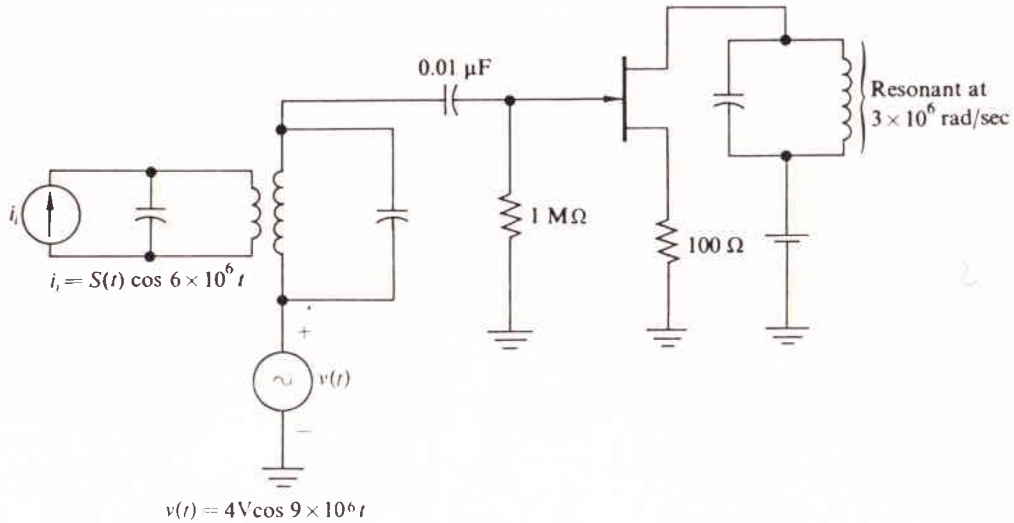

0.01 μF
Resonant at
3 × 10⁶ rad/sec
i_i
1 MΩ
100 Ω
$i_i = S(t) \cos 6 \times 10^6 t$
+
$v(t)$
−
$v(t) = 4\text{V} \cos 9 \times 10^6 t$

Figure 7.P–1

7.5 Estimate the conversion transconductance of the mixer shown in Fig. 7.P–1. Use the same transistor parameters as in Problem 7.1. The oscillator peak voltage is 4 V.

7.6 Consider the differential-pair mixer shown in Fig. 7.4–4. How large can V_1 be without introducing more than 7% second-harmonic distortion in I_k (assuming a pure sine-wave drive from the $V_1 \cos\omega_0 t$ generator)? Why is it desirable to keep the second-harmonic component in I_k small? Why is it desirable that the voltage-current characteristic of the differential pair be symmetrical?

7.7 For the mixer circuits shown in Fig. 7.P–2, show that the equivalent linear loading the transistor places across the input-tuned circuit is given by

$$G_{RF} = \frac{v_{RF}}{i_{RF}} = \frac{g_{in}}{1 + \beta}\left[1 + \frac{\ln I_0(x)}{qV_\lambda/kT}\right],$$

where $V_\lambda = V_{EE} - V_0$, $x = qV_1/kT$, and $g_{in} = qI_{EQ}/kT$.

7.8 For the mixer circuit shown in Fig. 7.P–3, determine an expression for $v_o(t)$ for the case where $g(t) = (1\text{ mV})[1 + mf(t)]$, $\omega_s = 9 \times 10^7$ rad/sec, and the output-tuned circuit has sufficient bandwidth to pass the modulation. (Assume Q_1, Q_2, and Q_3 to be identical and that $\alpha \approx 1$.)

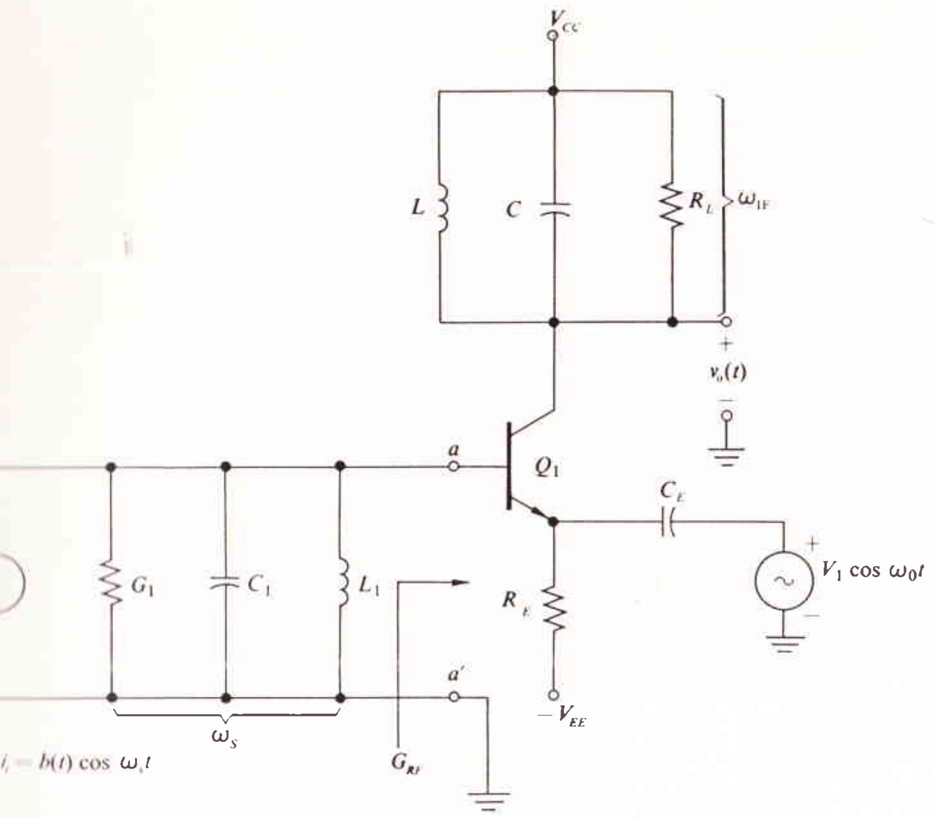


Figure 7.P–2

For the converter shown in Fig. 7.P–4, determine an expression for $v_o(t)$. Justify all assumptions. $\omega_s = 9 \times 10^7$.

For the "front end" shown in Fig. 7.P–5, determine an expression for $v_o(t)$. Justify all assumptions and do not neglect the loading of L_2 by Q_1 (cf. Problem 7.7). Assume that L_0C_0 passes the modulation. At the frequencies of operation, transistor reactances are negligible.

A nonlinear device has an input voltage–output current transfer characteristic given by

$$i_0 = a_0 + a_1 v_1 + a_2 v_1^2 + a_4 v_1^4.$$

If $v_1 = v_s(t) \cos \omega_s t + V_1 \cos \omega_0 t$, where $|v_s(t)| \ll V_1$, show that the voltage-controlled transconductance of such a device for $v_s \cos \omega_s t$ is given by

$$g_m(V_1 \cos \omega_0 t) = a_1 + 2a_2 V_1 \cos \omega_0 t + 4a_4 (V_1 \cos \omega_0 t)^3.$$

Also show that the conversion transconductance between ω_s and $\omega_0 - \omega_s$ is given by

$$G = a_2 V_1 + 3/2 a_4 V_1^3.$$

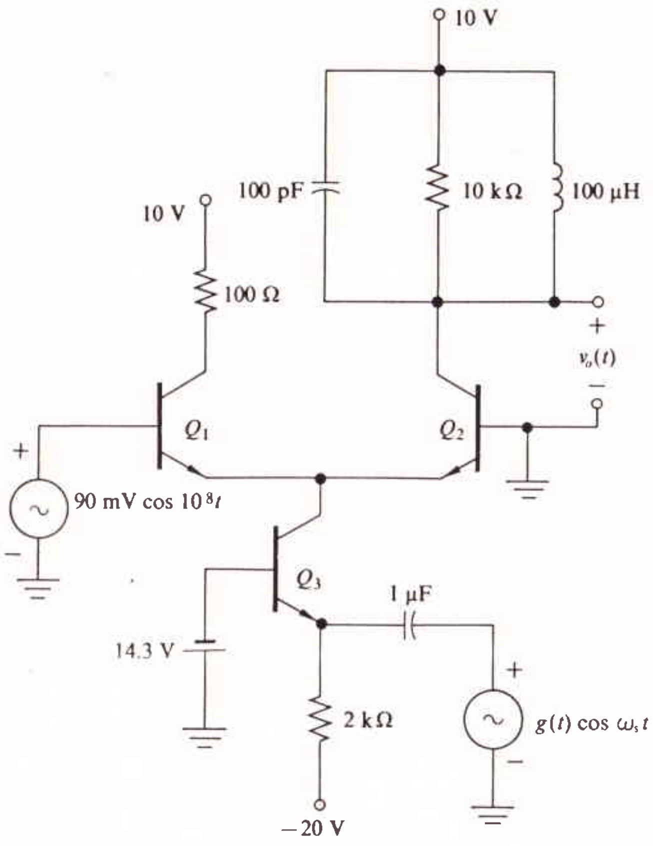


Figure 7.P–3

7.12 It is possible to choose an IF frequency at either $\omega_0 + \omega_s$ or $|\omega_0 - \omega_s|$. Explain in terms of the local oscillator tuning complexity why $|\omega_0 - \omega_s|$ is almost always chosen for the intermediate frequency.

7.13 Show that for the AVC circuit of Fig. 7.P–6 the envelope detector output is given by

$$\frac{-V_1 A_0}{1 - \dfrac{A_0 V_1}{V_{co}}}[1 + mf(t)].$$

Also show that the envelope detector output approaches

$$V_{co}[1 + mf(t)]$$

for $A_0 V_1 \gg |V_{co}|$. In this case the output level of the envelope detector is insensitive to input variations.

7.14 With the assumption that the modulation is not fed around the AVC loop of Problem 7.13, show that the poles of the closed-loop system are given by the roots of

$$V_{co} - V_1 A_0 H_{IFL}(p) H_L(p) = 0,$$

where $H_{\text{IF}\,L}(p)$ is the low-pass equivalent transfer function of the IF filter and $H_L(p)$ is the transfer function of the low-pass filter. Using a three-pole IF filter, show by means of a root locus plot how oscillations are possible as A_0V_1/V_{co} is increased to provide an output level which is insensitive to input-level variations. Show also how increasing the RC time constant of $H_L(p)$ reduces the possibility of oscillations at the expense of a "sluggish" AVC loop.

15 A simplified version of a 10.7 MHz IF amplifier is shown in Fig. 7.P–7. The transistor is a 2N918 biased at 2 mA with $V_{CE} = 10$ V. At the collector, the admittance seen by the transistor consists of a parallel combination of 8 kΩ, 3 μH, and C_v, where C_v is to be adjusted, in conjunction with C_{22} of the transistor, so as to produce resonance at 67 × 10^6 rad/sec.

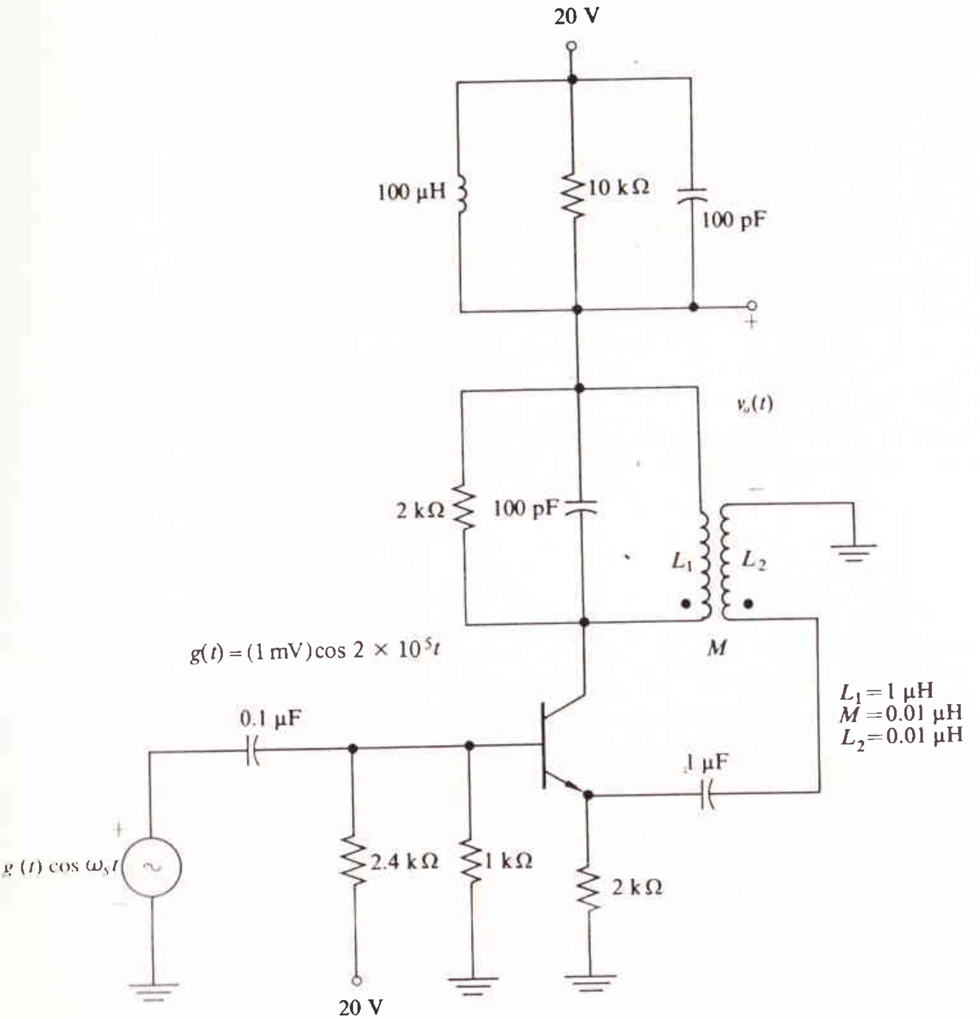


Figure 7.P–4

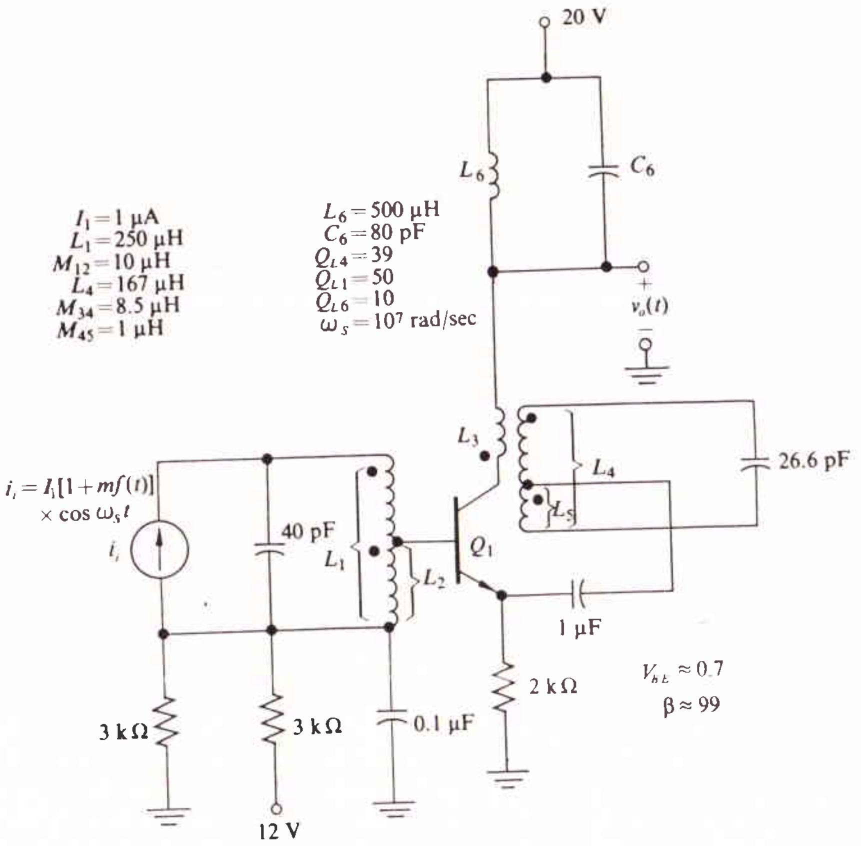


Figure 7.P–5

Choose values of input conductance and susceptance (to be presented at the base tap) so that the overall base admittance will be real at $\omega = 67 \times 10^6$ rad/sec and so that the alignability factor at this point will be 0.15. Do the numerical values that result seem realistic? Explain.

7.16 Redistribute the loss between the input and output circuits of Problem 7.15 so that the bandwidths are roughly equal and neither Q is larger than 30, while at the same time keeping $k < 0.15$. Sketch the shape of the bandpass (to the ±3 dB points) that would be seen at the base circuit under this new arrangement. (Assume a current source for the signal supply at the base terminals.)

7.17 Reexamine Problem 7.16 assuming that a 0.9 pF "neutralizing" capacitor is properly connected (via a unity voltage gain arrangement) across the transistor.

7.18 For the circuit of Problem 7.17, find the effect on the base circuit passband of varying the dc emitter current to 1 mA and then to 3 mA, while leaving all circuit components at the values found for the 2 mA case.

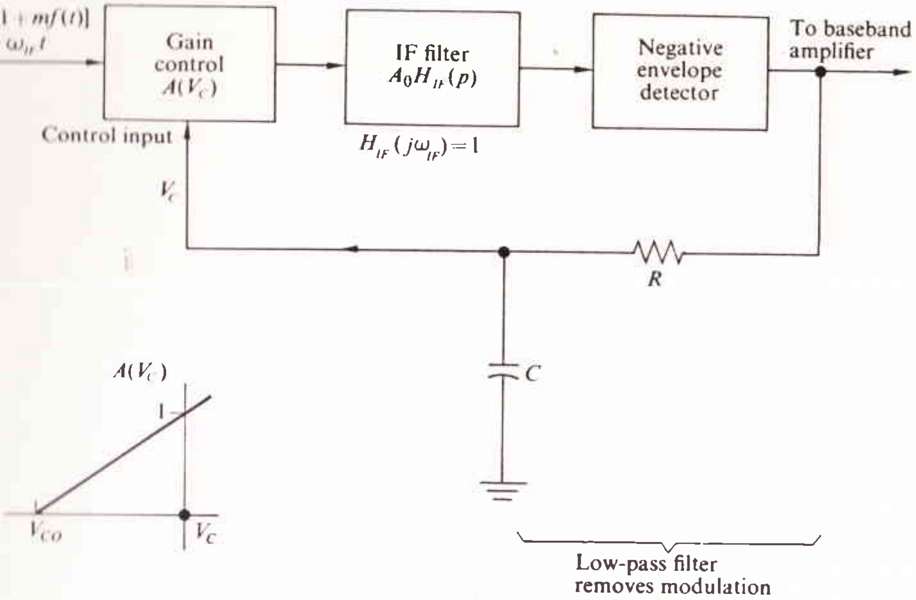

1 + mf(t)]
$\omega_{IF} t$
Gain control $A(V_c)$
Control input
V_c
IF filter $A_0 H_{IF}(p)$
$H_{IF}(j\omega_{IF}) = 1$
Negative envelope detector
To baseband amplifier
R
C
$A(V_c)$
1
V_{CO}
V_C
Low-pass filter removes modulation

Figure 7.P–6

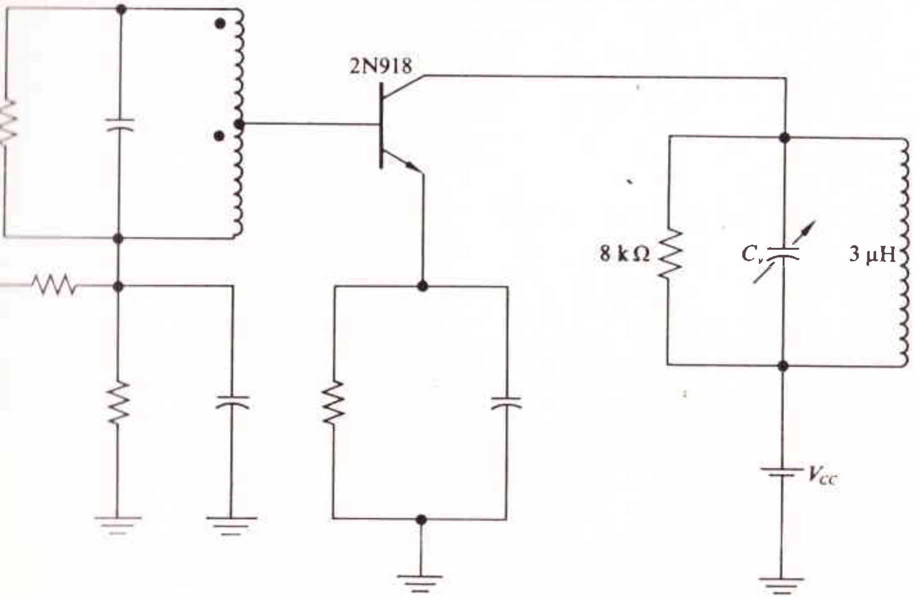


Figure 7.P–7

APPENDIX TO CHAPTER 7

COMPARISONS OF y-PARAMETERS FOR BIPOLAR TRANSISTORS: SINGLE-ENDED, DIFFERENTIAL-PAIR, AND CASCODE

If one assumes that identical transistors are used in all three configurations, if one neglects various parasitic reactances and resistances, and if one assumes simplified models for the transistors, then one can calculate the y-parameters for each case in a relatively straightforward fashion. The purpose of this appendix is not to do this algebra, but merely to point out the approximate comparative values to be expected from the three cases. Since the y-parameters are a function of frequency and since we are only trying to give order-of-magnitude results, one must not attempt to apply these results to precise comparisons between circuits.

If we assume equal power-supply currents for all three configurations, then the differential-pair transistors each have only half the dc current of all other transistors, hence their values for r_E will be doubled and their values for g_m will be halved. In addition, though their values for alpha or beta may be changed, we shall ignore such changes.

With these assumptions and the realization that from the small-signal viewpoint the differential pair consists of an emitter follower driving a common base stage, while the cascode connection consists of a common emitter stage driving a common base stage, we can make the comparisons shown in Table 7.A–1.

Table 7.A–1

	Differential pair compared to common emitter stage with twice the dc bias current of each transistor of the differential pair	Cascode connection compared to common emitter stage with the same dc bias current
y_{11}	1/4	1
y_{12}	1/30 to 1/200	1/200 to 1/2000
y_{21}	1/4	1
y_{22}†	1 to 1/3	1 to 1/3

† For $\omega C_{B'E} r_{BB} \ll 1$, $y_{22} \approx pC_{B'C}(1 + g_m r_{BB'})$, while for the two configurations with common base output stages it is approximately $pC_{B'C}$. (Ideally the cascode case should have an even smaller y_{22}; in practice, parasitic terms tend to keep it from being much smaller.)

Inspection of the results in Table 7.6–2 will indicate that the results shown in Table 7.A–1 are generally similar to those reported earlier. The y_{11} results for the

ferential pair and the cascode connection show the expected 4/1 differential;
wever y_{11} for the single-ended stage (not the same transistor in this case) is con-
erably smaller than the cascode value. This implies that the 2N918 used for the
gle-ended example is a "better" high-frequency transistor than those used in the
3005 in that it has a smaller total input capacitance (by a factor of about four)
d a higher beta (by a factor of about three). These conclusions follow from the fact
at, over the range of frequencies being considered, the common emitter stage has
input admittance

$$y_{11} \approx g_{B'C} + p(C_{B'C} + C_{B'E}) \qquad \text{or} \qquad y_{11} \approx g_{\pi} + p(C_{B'C} + C_{\pi}),$$

pending on one's choice of notation.

CHAPTER 8

AMPLITUDE MODULATION

8.1 AMPLITUDE MODULATION SIGNALS

If the amplitude, frequency, or phase of a high-frequency sinusoid is forced to vary in proportion to a desired low-frequency signal $f(t)$, a modulated signal is generated whose frequency spectrum is concentrated in the vicinity of the frequency of the unmodulated high-frequency sinusoid. The modulated signal, unlike the modulation $f(t)$, may be frequency-division multiplexed with other similarly modulated signals at different center frequencies (cf. Chapters 3 and 7); it can also be transmitted efficiently by antennas not nearly so large as those which would be required to transmit $f(t)$ directly. Specifically, if $f(t)$ were a single tone at a frequency of 1 kHz, a $\frac{1}{4}$ wavelength antenna to transmit $f(t)$ into the atmosphere would be 75 km long, whereas if $f(t)$ modulated a 100 MHz sinusoid the corresponding antenna length would be only $\frac{3}{4}$ m.

In this chapter we consider circuits that can amplitude-modulate a sinusoidal carrier; in Chapter 11 we shall focus our attention on circuits which modulate the frequency and phase of a similar carrier. Then in Chapters 10 and 12 respectively, we examine the techniques by which amplitude- and frequency-modulated carriers are processed to recover $f(t)$ at the receiver.

A typical amplitude-modulated carrier has the form

$$v(t) = A[1 + mf(t)]\cos(\omega_0 t + \theta_0) = g(t)\cos(\omega_0 t + \theta_0), \tag{8.1-1}$$

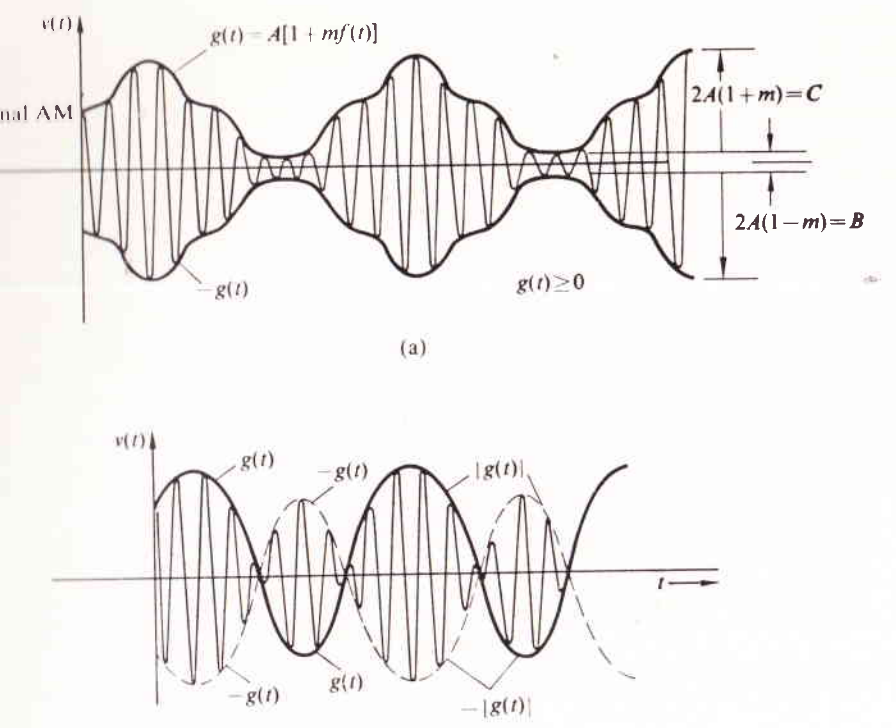
where A is the carrier amplitude, m is the modulation index, $f(t)$ is a signal which is proportional to the modulation information and which has the properties

$$|f(t)|_{\max} = 1 \qquad \text{and} \qquad \overline{f(t)} = 0, \tag{8.1-2}$$

ω_0 is the carrier frequency, θ_0 is the carrier phase angle (which is usually chosen as zero without loss of generality), and $g(t) = A[1 + mf(t)]$ is the envelope function of the AM wave. For the case where $g(t) \geq 0$ for all time, or equivalently $m \leq 1$ (which we refer to as *normal* AM), $g(t)$ forms the upper envelope of $v(t)$ while $-g(t)$ forms the lower envelope of $v(t)$, as illustrated in Fig. 8.1-1. For normal AM the modulation index can be obtained directly from the waveform of $v(t)$ by means of the relationship

$$m = \frac{C - B}{C + B}, \tag{8.1-3}$$

where C is the maximum peak-to-peak value of $v(t)$ and B is the minimum peak-to-peak value of $v(t)$. On the other hand, if $g(t)$ does not remain positive for all time,



(a)

$g(t)$ not always greater than zero

(b)

.1–1 Typical AM signals for which (a) $g(t) \geq 0$ for all time and (b) $g(t)$ does not remain
r than zero for all time.

ınd $-|g(t)|$ form the upper and lower envelopes of $v(t)$ respectively, as illustrated
, 8.1–1. For this case m can no longer be obtained from Eq. (8.1–3). As we shall
Chapter 10, only normal AM can be demodulated by simple envelope demodu-
s (or detectors). In all other cases synchronous detection, which is much more
plicated to implement, must be employed. Consequently almost all commercial
tations transmit normal AM to keep the receiver as simple as possible.

ıs derived in Chapter 3, the Fourier transform $V(\omega)$ of $v(t)$ has the form

$$V(\omega) = \tfrac{1}{2}[G(\omega + \omega_0) + G(\omega - \omega_0)], \tag{8.1–4}$$

e $G(\omega) = A[2\pi\delta(\omega) + mF(\omega)]$ is the Fourier transform of $g(t)$ and $F(\omega)$ is the
ier transform of $f(t)$. A sketch of $|V(\omega)|$ vs. ω is again presented in Fig. 8.1–2 for
ıse where $F(\omega)$ and in turn $G(\omega)$ are band-limited to $\omega_m < \omega_0$. As illustrated in
igure, the portion of the spectrum lying above ω_0 is referred to as the upper
and of $V(\omega)$, whereas the portion of the spectrum lying below ω_0 (and above
0) is referred to as the lower sideband of $V(\omega)$.

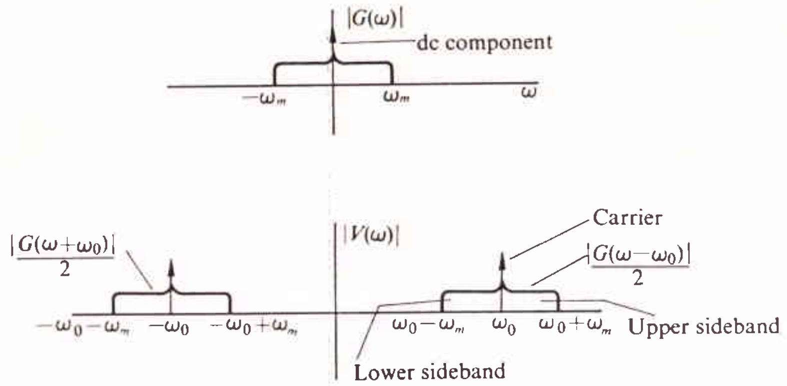


Fig. 8.1–2 Frequency spectrum of $v(t)$.

If a normal AM signal $v(t)$ is placed across a 1 Ω resistor, the power delivered to the resistor is given by†

$$\begin{aligned} P_{1\Omega} &= \overline{A^2 \cos^2 \omega_0 t[1 + 2mf(t) + m^2 f^2(t)]} \\ &\doteq \overline{A^2 \cos^2 \omega_0 t}[1 + \overline{m^2 f^2(t)}] \\ &= \frac{A^2}{2} + \frac{A^2 m^2}{2}\overline{f^2(t)} \\ &= P_c + P_m, \end{aligned} \tag{8.1–5}$$

where $P_c = A^2/2$ is the carrier power, $P_m = (A^2 m^2/2)\overline{f^2(t)}$ is the modulation power, and $\overline{f^2}$ denotes the time average of f^2 or equivalently the mean square value of $f(t)$. Since $|f(t)| \leq 1$, then $\overline{f^2(t)} \leq 1$; hence the carrier power is greater than or equal to the modulation power (or the sideband power) even with $m = 1$ (100% modulation). For example, if $f(t) = \cos \omega_m t$ and $m = 1$, then $\overline{f^2(t)} = \frac{1}{2}$ and $P_c = 2P_m$. If $f(t) = \cos \omega_m t$ and $m = \frac{1}{2}$, the modulation power has decreased to one-eighth of the carrier power. For efficient transmission of normal AM, it is clear that we want the modulation index to be as close to unity as possible.

Because of the large amount of carrier power, which carries no information, that must be transmitted in a normal AM signal, the carrier is sometimes removed (or not generated in the first place) before transmitting the AM signal. The result is a suppressed carrier AM signal of the form

$$\begin{aligned} v(t) &= \overbrace{Am}^{A'} f(t) \cos \omega_0 t \\ &= g(t) \cos \omega_0(t), \end{aligned} \tag{8.1–6}$$

where $g(t) = A'f(t)$ and $\overline{g(t)} = 0$. In the time domain a suppressed carrier signal has the form shown in Fig. 8.1–1(b), whereas in the frequency domain the spectrum is lentical with that shown in Fig. 8.1–2 except that the impulses at $\pm\omega_0$ are absent. he price that must be paid for the reduced power at the transmitter is increased eceiver complexity.

---

† In evaluating $P_{1\Omega}$ we use the relationships $\overline{A + B} = \bar{A} + \bar{B}$ and $\overline{AB} = \overline{A}\,\overline{B}$. The second relationship is valid only if A and B are independent as $f(t)$ and $\cos \omega_0 t$ are.

Still a further saving in transmitted power, as well as a reduction in the frequency pectrum, can be achieved by removing the upper or lower sideband of a suppressed arrier AM signal to form a *single-sideband* AM signal. A filtering technique, to form single-sideband signal, illustrated in Fig. 8.1–3, is justified by the fact that all the nformation necessary to reconstruct $g(t)$ is completely contained in the positive (or egative) portion of $G(\omega)$. If $g(t)$ is real (which it always is), then $G(\omega) = G^*(-\omega)$.

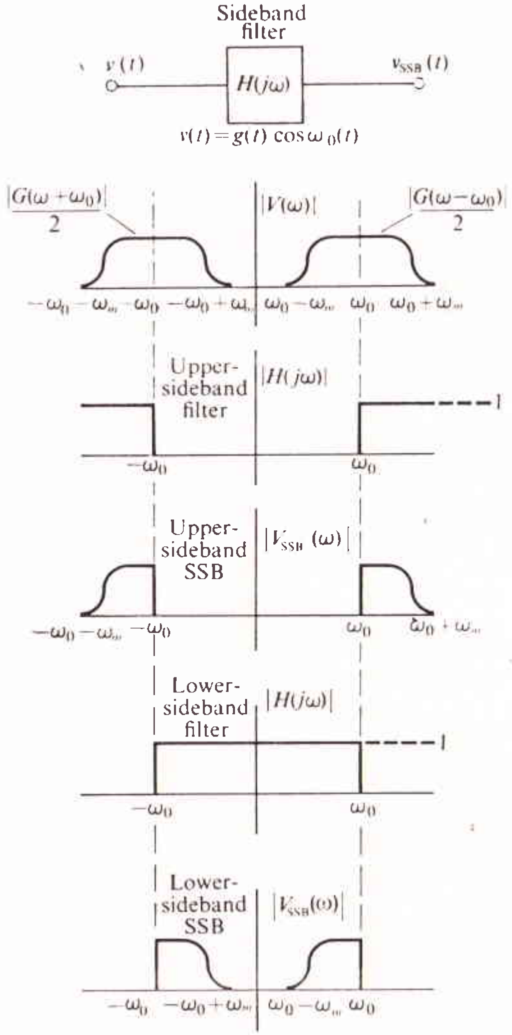


Fig. 8.1–3 Generation of upper and lower single-sideband signals.

Therefore, the positive and negative portions of $G(\omega)$ contain identical information; i.e., the negative spectrum can be completely constructed if the positive spectrum is known.

To evaluate an expression for the single-sideband signal v_{SSB} in the time domain, we obtain the inverse Fourier transform of $V_{SSB}(\omega)$ which for upper-sideband SSB is given by

$$V_{SSB}(\omega) = \begin{cases} \dfrac{G(\omega+\omega_0)}{2} + \dfrac{G(\omega-\omega_0)}{2}, & |\omega| > \omega_0, \\ 0, & |\omega| \le \omega_0. \end{cases} \tag{8.1–7}$$

Assuming that $g(t)$ is band-limited to $\omega_m < \omega_0$, we obtain

$$\begin{aligned} v_{SSB}(t) &= \frac{1}{2\pi}\int_{-\infty}^{-\omega_0} \frac{G(\omega+\omega_0)}{2} e^{j\omega t}\,d\omega + \frac{1}{2\pi}\int_{\omega_0}^{\infty} \frac{G(\omega-\omega_0)}{2} e^{j\omega t}\,d\omega \\ &= e^{-j\omega_0 t}\left[\frac{1}{2\pi}\int_{-\infty}^{0} \frac{G(\omega)}{2} e^{j\omega t}\,d\omega\right] + e^{j\omega_0 t}\left[\frac{1}{2\pi}\int_{0}^{\infty} \frac{G(\omega)}{2} e^{j\omega t}\,d\omega\right] \\ &= \frac{\cos\omega_0 t}{2}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} G(\omega) e^{j\omega t}\,d\omega\right] \\ &\quad - \frac{\sin\omega_0 t}{2}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} G(\omega)(-j\,\text{sgn}\,\omega) e^{j\omega t}\,d\omega\right] \\ &= \frac{g(t)}{2}\cos\omega_0 t - \frac{\hat{g}(t)}{2}\sin\omega_0 t, \end{aligned} \tag{8.1–8}$$

where $\hat{g}(t) \equiv \mathscr{F}^{-1}[G(\omega)(-j\,\text{sgn}\,\omega)]$ is referred to as the Hilbert transform of $g(t)$ and sgn ω is given by

$$\text{sgn}\,\omega = \begin{cases} 1, & \omega > 0, \\ 0, & \omega = 0, \\ -1, & \omega < 0. \end{cases} \tag{8.1–9}$$

In a similar fashion, lower-sideband SSB can be shown to have the form

$$v_{SSB}(t) = \frac{g(t)}{2}\cos\omega_0 t + \frac{\hat{g}(t)}{2}\sin\omega_0 t. \tag{8.1–10}$$

Note that $\hat{g}(t)$ is obtained by placing $g(t)$ through a linear filter having the transfer function $H_{HT}(j\omega) = -j\,\text{sgn}\,\omega$. The filter $H_{HT}(j\omega)$, which is referred to as a Hilbert transformer, has the magnitude and phase characteristic shown in Fig. 8.1–4. Clearly by shifting all the positive frequency components of $g(t)$ by $-\pi/2$ we obtain $\hat{g}(t)$.

In addition to characterizing SSB in the time domain, Eqs. (8.1–8) and (8.1–10) suggest an alternative method of generating SSB which is shown in block diagram

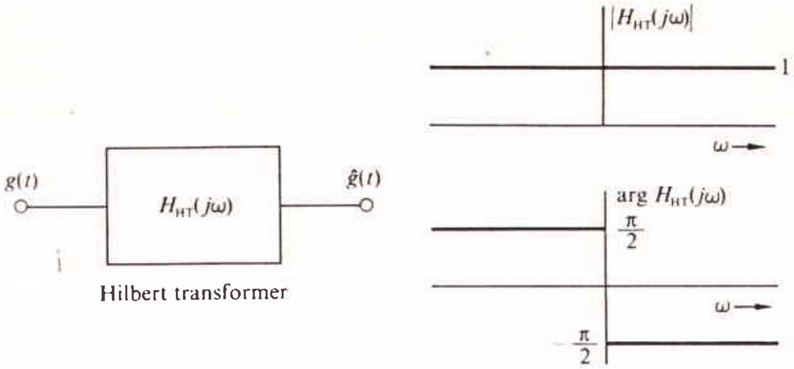


Fig. 8.1–4 Plot of magnitude and phase of the Hilbert transformer.

rm in Fig. 8.1–5. Here $g(t)$ is suppressed-carrier-modulated (or multiplied) by
s $\omega_0 t$ while $\hat{g}(t)$ is suppressed-carrier-modulated by $-\sin \omega_0 t$. The resultant sum
r difference) yields Eq. (8.1–8) or Eq. (8.1–10).

Although the block diagram of Fig. 8.1–5 appears to be a very reasonable method
f generating SSB, it encounters the severe practical problem that the Hilbert trans-
rmer is a nonrealizable filter (its impulse response $h_{HT}(t) = 1/\pi t$ is noncausal).
onsequently its response may at best be approximated only over a reasonably
arrow band of frequencies. In addition, such approximations usually produce a
hase of the form

$$\arg H_{HT}(j\omega) \approx -\frac{\pi}{2} + \theta(\omega),$$

hich must be compensated for by placing $g(t)$ through a filter with constant magni-
ude and with a phase of $\theta(\omega)$ before multiplying it by $\cos \omega_0 t$. In general, no really
atisfactory SSB modulators of the form shown in Fig. 8.1–5 have been produced.

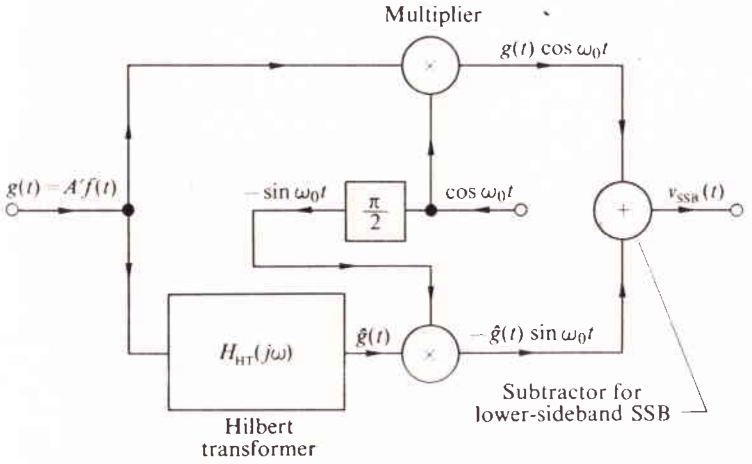


Fig. 8.1–5 Alternative method of generating a single sideband.

Even direct sideband filtering of suppressed carrier AM to produce SSB has its practical limitations. It is a difficult matter to design a sideband filter which cuts off sufficiently rapidly to attenuate one sideband while not distorting the other sideband either in magnitude or in phase. For SSB voice modulation, mechanical, crystal, or ceramic bandpass filters are usually employed to remove the undesired sideband (See Fig. 7.6–1 for an example). Even though these filters have considerable ripple in magnitude as well as significant nonlinearity in phase in the passband, their effect on the intelligibility of average speech is negligible. For other forms of modulation the sideband filter usually has to be hand-tailored to the modulation to minimize distortion. In all cases, however, the sharp cutoff required of the bandpass sideband filter is possible only if the filter center frequency is not too high (the required Q of the tuned circuits in the filter to maintain a fixed BW is directly proportional to ω_0). Consequently, in almost all SSB transmitters, the information is suppressed-carrier-modulated at a reasonably low carrier frequency (50 kHz to 500 kHz for voice modulation) where sideband filtering is accomplished, and then the resultant SSB signal is heterodyned to the desired carrier frequency ω_0.

In addition to the complications that SSB creates at the transmitter, its demodulation is possible only by synchronous detection, which requires a reference oscillator at the radian frequency ω_0. Since it is impossible to derive this frequency from the SSB signal itself, a small pilot carrier is usually transmitted along with the SSB signal to provide the reference at the receiver. The demodulation of SSB will be pursued in more detail in Chapter 10.

In the subsequent sections of this chapter we shall consider the theoretical methods by which amplitude modulation (or multiplication of two signals) may be accomplished. We shall then examine some practical circuits which implement the theoretical methods.

8.2 AMPLITUDE MODULATION TECHNIQUES

In this section we investigate the theoretical methods by which we can multiply or modulate $\cos \omega_0 t$ by $g(t)$ to obtain the AM signal

$$\begin{aligned} v(t) &= g(t) \cos \omega_0 t \\ &= A[1 + mf(t)] \cos \omega_0 t. \qquad (8.2\text{–}1) \end{aligned}$$

In general, there are four basic methods by which amplitude modulation can be accomplished:

a) analog multiplication,
b) chopper modulation,
c) nonlinear device modulation, and
d) direct tuned-circuit modulation.

As we shall see, all these methods can be employed to generate normal AM, whereas only direct analog multiplication and chopper modulation can be employed to generate suppressed carrier AM. In addition, with the exception of the direct tuned-circuit modulator, modulation is accomplished at low power levels and amplified

INDEX

Abramowitz, M., 641
AFC, 590
AGC, 308, 331–336, 486, 530
Alignability, 324–327
AM generators: balanced, 382–384
  differential amplifier, 374–375
  diode bridge, 376–379, 382–384
  FET, 380–381
  integrated circuit, 395–396
  square law, 384–387
  tuned circuit, 387–392
  See also Diode bridge, FET, Chopper modulation, Multiplier, *and/or* Modulation
AM methods: analog, 354–355
  chopper, 355–358
  direct, 361–362
  high level, 447–450
  nonlinear, 359–360
AM on FM signals, 589–590
AM receiver, superheterodyne, 478
AM signals: in capacitively coupled circuits, 168–169, 195–199
  in narrowband filters, 72, 76, 88–89
  normal, 347
  single sideband, 350–352, 457–461, 464–466
  suppressed carrier, 349, 357
Amplification, AM, 405–406
Amplifiers: cascode, 324–326
  class A, 401–405
  class AB, 401
  class B, broadband, 401, 421–425
  class B "linear," 92, 405–408
  class C "linear," 408–410
  class C–RF, 401, 410–415
  class D, broadband, 426–432
  class D, narrowband, 401, 415–421
  differential pair, 114–119, 324–326
  feedback, 206, 475–478
  FET, 98–104, 131–136
  gain controlled, 214, 335–336
  maximum signal without distortion, 4, 100, 110
  narrowband-tuned, 6–8, 314–331, 405–421
  power, 401–455
  small signal range, 4–5, 110, 118
  transistor, 1–8, 104–113
  tuned, 6–8, 314–331, 405–421
Amplitude control, 212–215, 265–266, 308, 331–336
Amplitude limiting, 7–8, 110, 188, 571–577
  nonlinear, 207
Amplitude modulation; *see specific headings under* AM
Amplitude sensitivity, function, 230–231, 264–265
Amplitude stabilization: FM, 530–532
  oscillators, 212–215, 262–266
Analog computer, 528
Angelo, E. J., Jr., 13
Armstrong, E. H., 560
Astable multivibrator, 553–556
Attenuator, nonlinear, 213, 265–266
Auto transformer, 49–51
Automatic frequency control (AFC), 590
Automatic gain control; *see* AGC
AVC; *see* AGC
Average envelope detection, 461–466
  in FM detectors, 582–584, 587
Average envelope detector, 468–478
  dual transistor, 473–475
  operational amplifier, 475–478
Avins, J., 613

ss B—see Section 4.2 and 9.2) to the desired output level. The direct tuned-circuit lulator directly modulates the amplitude of a high-power carrier which has been lified by more efficient class C amplifiers to the desired level.

log Modulation

log modulation (or multiplication) is accomplished in any device whose output)] is directly proportional to two inputs [$v_1(t)$ and $v_2(t)$], that is,

$$v_o(t) = Kv_1(t)v_2(t). \tag{8.2–2}$$

rly, if $v_1(t) = \cos\omega_0 t$ and $v_2(t) = g(t)$, then $v_o(t) = Kg(t)\cos\omega_0 t$, which is the red AM wave. In such devices no theoretical limitations exist; however, practical ce limitations usually impose limits on both the amplitude and frequency of), $v_2(t)$, and $v_o(t)$ in order to maintain the validity of Eq. (8.2–2).

Analog modulation can also be accomplished with two square-law devices as

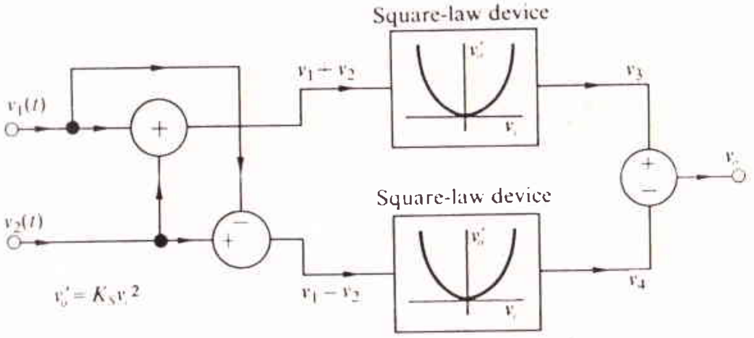


Fig. 8.2–1 Analog modulator constructed with two square-law devices.

wn in Fig. 8.2–1. In this system the output of the top square-law device, v_3, is en by

$$v_3 = K_S(v_1^2 + 2v_1v_2 + v_2^2); \tag{8.2–3}$$

output of the bottom square-law device, v_4, is given by

$$v_4 = K_S(v_1^2 - 2v_1v_2 + v_2^2). \tag{8.2–4}$$

system output $v_o = v_3 - v_4$ is thus given by

$$v_o(t) = 4K_S v_1(t)v_2(t), \tag{8.2–5}$$

ich is the form of the output of an analog modulator.

If the square-law devices are "half square law" rather than full square law, i.e., if

$$v_o' = \begin{cases} K_S v_i^2, & v_i > 0, \\ 0, & v_i \le 0, \end{cases} \tag{8.2–6}$$

en $v_1 + v_2$ and $v_1 - v_2$ must be constrained to be greater than zero to produce the

desired output. Thus if $v_1 = A[1 + mf(t)]$ and $v_2 = V_1\cos\omega_0 t$, where $f(t)$ is the modulation information (cf. Eq. 8.1–2), then

$$v_o(t) = \overbrace{4K_S V_1 A}^{A'}[1 + mf(t)]\cos\omega_0 t, \tag{8.2–7}$$

provided that $A(1 - m) - V_1 \ge 0$ or equivalently

$$m \le 1 - \frac{V_1}{A}. \tag{8.2–8}$$

Since $V_1 > 0$, we observe that, with half-square-law devices in the circuit of Fig. 8.2–1, the modulation index is constrained to be less than unity and thus the circuit clearly cannot be used to generate suppressed carrier AM.

Even though the circuit does not produce a 100% modulated AM wave (which is desirable for efficient transmission), the modulation index at the output can be increased by subtracting some of the excess carrier from the undermodulated signal. This technique is referred to as *carrier cancellation*. For example, if we subtract $D\cos\omega_0 t$ from $v(t) = A[1 + mf(t)]\cos\omega_0 t$, we obtain $v_o(t)$ in the form

$$v_o(t) = \overbrace{(A - D)}^{A'}\left[1 + \overbrace{\frac{Am}{A - D}}^{m'} f(t)\cos\omega_0 t\right], \tag{8.2–9}$$

from which it is apparent that we can increase the resultant modulation index m' to unity by choosing D such that $(A - D)/A = m$ or $D/A = 1 - m$.

Chopper Modulation

Chopper modulation is accomplished by chopping $g(t)$ at the carrier frequency rate and placing the resultant signal through a bandpass filter centered at the carrier frequency. The basic skeleton circuit of the chopper modulator is shown in Fig. 8.2–2, in which the switch, which is controlled by $\cos\omega_0 t$, remains open for $A\cos\omega_0 t \ge 0$ and closed for $A\cos\omega_0 t < 0$. To demonstrate that this circuit accomplishes amplitude modulation, we first write $v_a(t)$ in the form

$$v_a(t) = g(t)S(t), \tag{8.2–10}$$

where $S(t)$ is a switching function having the properties

$$S(t) = \begin{cases} 1, & \cos\omega_0 t \ge 0, \\ 0, & \cos\omega_0 t < 0. \end{cases} \tag{8.2–11}$$

Clearly, then, $S(t)$ is a square wave with unity amplitude which may be expanded in a Fourier series of the form

$$S(t) = \frac{1}{2} + \frac{2}{\pi}\cos\omega_0 t - \frac{2}{3\pi}\cos 3\omega_0 t + \cdots \tag{8.2–12}$$

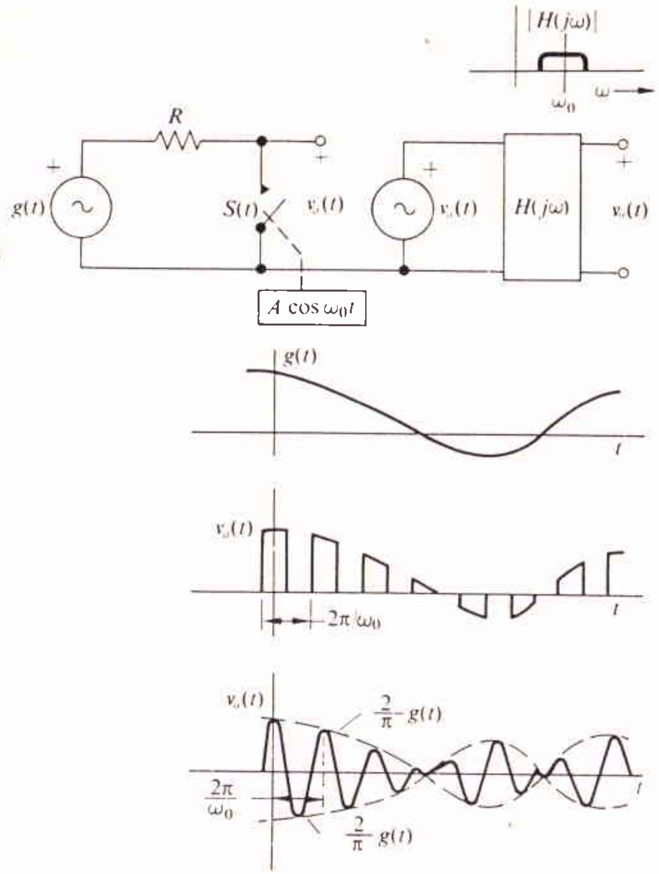


Fig. 8.2–2 Chopper modulator.

yield

$$v_a(t) = \frac{g(t)}{2} + \frac{2g(t)}{\pi} \cos \omega_0 t - \frac{2g(t)}{3\pi} \cos 3\omega_0 t + \cdots \tag{8.2–13}$$

rom Eq. (8.2–13) we note that $v_a(t)$ is the superposition of AM waves centered at $_0$, $3\omega_0$, $5\omega_0, \ldots$. If the bandpass filter $H(j\omega)$ attenuates the low-frequency comnents of $v_a(t)$ as well as the AM components of $v_a(t)$ in the vicinity of $3\omega_0, 5\omega_0, \ldots,$ en the output $v_o(t)$ is given by

$$v_o(t) = \left[\frac{2g(t)}{\pi} * h_L(t)\right] \cos [\omega_0 t + \theta(\omega_0)], \tag{8.2–14}$$

here $h_L(t)$ is the impulse response of the low-pass equivalent of the bandpass filter d $\theta(\omega_0)$ is the plane angle of $H(j\omega)$ at $\omega = \omega_0$. If the low-pass equivalent filter is

---

flat over the band of frequencies occupied by $g(t)$, then $v_o(t)$ simplifies to the desired form

$$v_o(t) = \frac{2g(t)}{\pi} H_L(0) \cos [\omega_0 t + \theta(\omega_0)], \tag{8.2–15}$$

where $H_L(j\omega)$ is the Fourier transform of $h_L(t)$.

Note that chopper modulation is possible only if the spectrum of the desired AM wave does not overlap the spectra of any of the other components of $v_a(t)$. A typical sketch of $|V_a(\omega)|$, where $V_a(\omega)$ is the Fourier transform of $v_a(t)$, is presented in Fig. 8.2–3 for the case where $g(t)$ is band-limited to ω_m. It is apparent from the figure that, unless

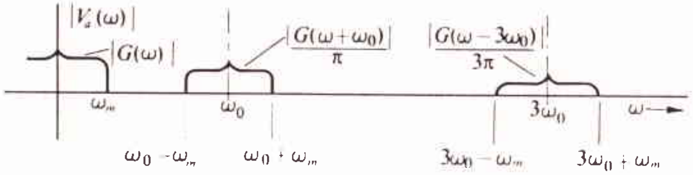


Fig. 8.2–3 Spectrum of $v_a(t)$.

$$\omega_m < \omega_0/2, \tag{8.2–16}$$

then the spectra overlap and chopper modulation is impossible. In addition, the closer ω_m is to $\omega_0/2$, the more complex the bandpass filter must be to effect the frequency separation.

Note that no restriction has been placed on the modulation index of $g(t)$; therefore, normal AM with a modulation index of unity as well as suppressed carrier AM may be generated with the chopper modulator.

In order to relax the inequality of Eq. (8.2–16) and in turn make the design of the filter $H(j\omega)$ simpler, the single-pole voltage-controlled switch of Fig. 8.2–2 may be replaced by the double-pole voltage-controlled reversing switch shown in Fig. 8.2–4. The reversing switch has the effect of making $v_a(t)$ symmetrical about zero,

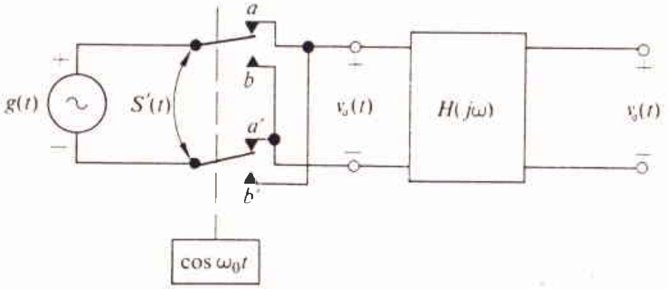


Fig. 8.2–4 Balanced chopper modulator.

thus eliminating the low-frequency component of $v_a(t)$. Consequently, with the chopper modulator of Fig. 8.2–4, modulation is possible if

$$\omega_m < \omega_0. \tag{8.2–17}$$

Equation (8.2–17) ensures that the spectra of the desired AM wave centered at ω_0 and the AM wave centered at $3\omega_0$ do not overlap.

On a rigorous basis we observe that for the balanced chopper modulator of Fig. 8.2–4

$$v_a(t) = g(t)S'(t) \tag{8.2–18}$$

where

$$S'(t) = \begin{cases} 1, & \cos\omega_0 t \geq 0, \\ -1, & \cos\omega_0 t < 0. \end{cases} \tag{8.2–19}$$

Since $S'(t)$ is a square wave with a peak-to-peak amplitude of 2 and *zero average value*, $S'(t)$ may be expanded in a Fourier series of the form

$$S'(t) = \frac{4}{\pi}\cos\omega_0 t - \frac{4}{3\pi}\cos 3\omega_0 t + \cdots. \tag{8.2–20}$$

Hence

$$v_a(t) = \frac{4g(t)}{\pi}\cos\omega_0 t - \frac{4g(t)}{3\pi}\cos 3\omega_0 t + \cdots, \tag{8.2–21}$$

which, as expected, has no low-frequency component. Equation (8.2–21) also indicates that the reversing switch permits us to obtain twice as much output from the modulator of Fig. 8.2–4 as from the modulator of Fig. 8.2–2.

In many practical cases the switch $S(t)$ in the chopper modulator of Fig. 8.2–2 has a resistance r in series with it which prevents complete attenuation of $g(t)$ when $S(t)$ is closed. Consequently, $v_a(t)$ is switched between $g(t)$ and $rg(t)/(r + R)$ in lieu of $g(t)$ and 0. This may be expressed mathematically as

$$v_a(t) = g(t)S(t) + \frac{g(t)[1 - S(t)]r}{r + R}. \tag{8.2–22}$$

It is apparent that the fundamental component of $v_a(t)$ is given by

$$\frac{2g(t)}{\pi}\left(1 - \frac{r}{r + R}\right)\cos\omega_0 t = \frac{2g(t)}{\pi}\frac{R}{R + r}\cos\omega_0 t, \tag{8.2–23}$$

and in turn $v_o(t)$ is given by

$$v_o(t) = \left[\frac{2g(t)}{\pi} * h_L(t)\right]\frac{R}{R + r}\cos[\omega_0 t + \theta(\omega_0)]. \tag{8.2–24}$$

Thus we see that the only effect of the series resistance r is to attenuate the output by the factor $R/(R + r)$. If $r = R$, the output level is reduced by a factor of 2.

Nonlinear Device Modulation

Nonlinear device modulation is accomplished by summing the modulation and the carrier, applying them to a nonlinear device, and then passing the device output through a bandpass filter centered at ω_0 to extract the desired AM signal. A block diagram of a nonlinear device modulator is shown in Fig. 8.2–5. As we shall see, the nonlinear device modulator has more restrictions for its proper operation than any of the previously considered modulators. First of all the nonlinear device must have no greater than a second-order (square-law) nonlinearity. Second, the maximum modulation frequency ω_m must be less than $\omega_0/3$; and third, if the nonlinear device contains a "half-square-law" term, 100% modulation or suppressed carrier modulation is not possible.

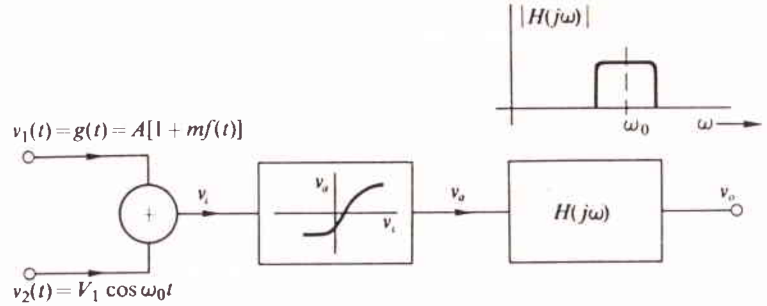


Fig. 8.2–5 Block diagram of nonlinear modulator.

To determine the reason for these restrictions as well as an expression for the output signal $v_o(t)$, let us first express the output of the nonlinear device $v_a(t)$ in the form of a MacLauren series,

$$v_a = a_0 + a_1 v_i + a_2 v_i^2 + a_3 v_i^3 + \cdots,$$

where

$$a_n = \frac{1}{n!}\frac{\partial v_a}{\partial v_i}\bigg|_{v_i = 0}. \tag{8.2–25}$$

With $v_i = v_1 + v_2$, v_a reduces to

$$\begin{aligned} v_a = a_0 &+ a_1(v_1 + v_2) + a_2(v_1^2 + 2v_1v_2 + v_2^2) \\ &+ a_3(v_1^3 + 3v_1^2v_2 + 3v_1v_2^2 + v_2^3) + \cdots. \end{aligned} \tag{8.2–26}$$

A little thought indicates that, if $v_2(t) = V_1\cos\omega_0 t$ and $v_1(t) = g(t)$, the following components of v_a (as well as a number of other components) have frequency spectra e vicinity of ω_0:

$$\begin{aligned} a_1 v_2 &= a_1 V_1 \cos \omega_0 t, \\ 2a_2 v_1 v_2 &= 2a_2 V_1 g(t) \cos \omega_0 t, \\ 3a_3 v_1^2 v_2 &= 3a_3 V_1 g^2(t) \cos \omega_0 t, \\ n a_n v_1^{n-1} v_2 &= n a_n V_1 g^{n-1}(t) \cos \omega_0 t. \end{aligned} \tag{8.2–27}$$

s apparent that any filter which extracts the desired term $2a_2V_1g(t)\cos\omega_0 t$ from
) also extracts the remaining terms of Eq. (8.2–27); however, the terms with the
fficient a_n for $n = 3, 4, 5, \ldots$ possess envelopes which are nonlinear functions of
Consequently for proper operation the MacLauren expansion of the non-
arity must have $a_n = 0$ for $n \geq 3$.

With this restriction, $v_a(t)$ simplifies to

$$v_a(t) = \overbrace{a_0 + \frac{a_2^2 V_1^2}{2} + a_1 g(t) + a_2 g^2(t)}^{\text{low-frequency term}} + \overbrace{V_1 \cos \omega_0 t[a_1 + 2a_2 g(t)]}^{\text{desired term}}$$
$$+ \overbrace{\frac{a_2}{2} V_1^2 \cos 2\omega_0 t}^{\text{second-harmonic term}}. \tag{8.2–28}$$

ce $g^2(t)$ has a spectrum which is band-limited to $2\omega_m$ if $g(t)$ is band-limited to
, it becomes apparent that the spectra of the desired component of $v_a(t)$ and the
-frequency component of $v_a(t)$ overlap unless

$$\frac{\omega_0}{3} > \omega_m. \tag{8.2–29}$$

However, if Eq. (8.2–29) is satisfied and the nonlinearity is "square law" ($a_n = 0$
$n \neq 2$), then $v_o(t)$ reduces to

$$v_o(t) = 2a_2 V_1[g(t) * h_L(t)] \cos [\omega_0 t + \theta(\omega_0)], \tag{8.2–30}$$

for the case where $H_L(j\omega)$ is flat over the band of frequencies occupied by $g(t) =$
$1 + mf(t)]$,

$$v_o(t) = \overbrace{2V_1 H_L(0) a_2 A}^{A'}[1 + mf(t)] \cos [\omega_0 t + \theta(\omega_0)], \tag{8.2–31}$$

ich is the desired AM wave.

If the nonlinear device is "half square law" rather than full square law, which is
ost always the case in practice, then the arguments leading to Eq. (8.2–8) apply
this case and the modulation index is limited to $m \leq 1 - V_1/A$. Hence again
ither 100% AM modulation nor suppressed carrier modulation is possible.

Direct Tuned-Circuit Modulation

Direct tuned-circuit modulation is effected by employing $g(t) = A[1 + mf(t)]$ to directly control the voltage across a parallel resonant circuit tuned to the carrier frequency and driven by a periodic current source. Figure 8.2–6 illustrates such a modulator. If the loaded Q_T of the tuned circuit is sufficiently high, it is apparent that $v_o(t)$ contains only a fundamental term. In addition, from Example 5.5–3 we recall that the peak value or envelope of $v_o(t)$ must be $g(t)$; consequently the high-level modulator produces the desired AM signal.

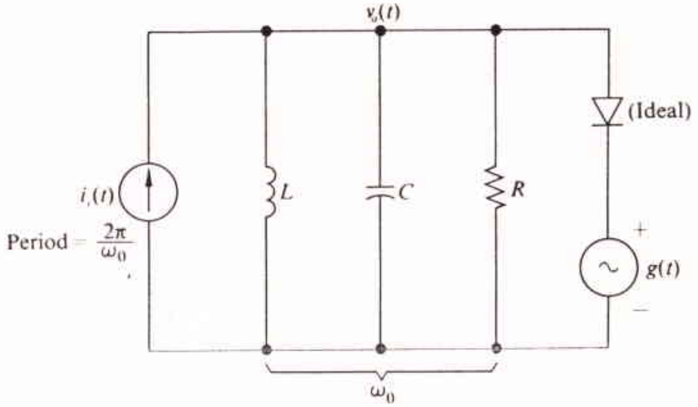


Fig. 8.2–6 Direct tuned-circuit modulator.

Like all the other modulators previously considered, the high-level modulator has a limit on the maximum modulation rate ω_m. If $g(t)$ increases too rapidly, the normal increase in the envelope of $v_o(t)$ cannot keep pace, the diode remains open, and the envelope of $v_o(t)$ is independent of $g(t)$. The resultant distortion is called *failure-to-follow distortion.* An exact relationship specifying the maximum permissible ω_m to prevent failure-to-follow distortion is obtained in Section 8.6.

In the direct circuit modulator, $g(t)$ is constrained to be greater than or equal to zero [a negative value of $g(t)$ would result in an infinite dc current through the diode]; therefore, suppressed carrier modulation is not possible even though normal AM with a modulation index of unity may be obtained. Since an exact expression for $v_o(t)$ as well as all the properties of the circuit of Fig. 8.2–6 are obtained in Section 8.6 and in Chapter 9, no further general analysis is attempted at this point.

Measurement of Envelope Distortion

When any of the modulators discussed in this section are implemented by practical circuits, the possibility exists that the circuits will introduce envelope distortion. For example, if $f(t)$ is the modulation information, then the modulator output might have the distorted form

$$v(t) = A[1 + mf_d(t)] \cos \omega_0 t, \tag{8.2–32}$$

where $f_d(t)$ is a nonlinear function of $f(t)$. In general, from direct observation of $v(t)$ the distortion is not apparent; however, if a Lissajous pattern of $v(t)$ vs. $Af(t)$ is formed on the face of an oscilloscope, a trapezoidal pattern results which makes any envelope nonlinearity clearly evident. A typical trapezoidal pattern is shown in Fig. 8.2–7.

If no envelope distortion exists, then $f_d(t) = f(t)$ and the upper and lower edges of the trapezoid are straight lines. Any departure of these edges from straight lines quickly makes evident the nature and magnitude of the nonlinear envelope distortion.

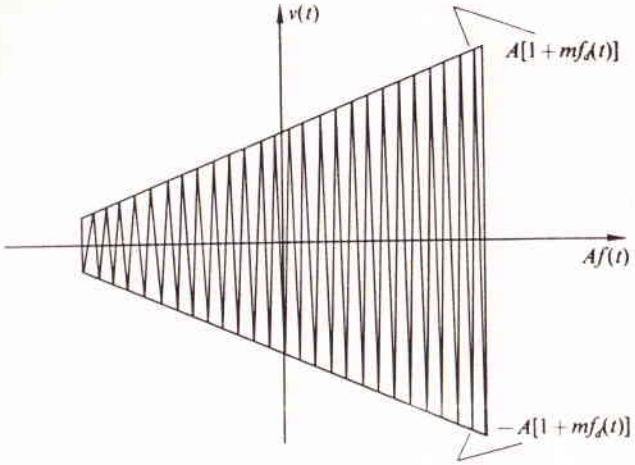


Fig. 8.2–7 Trapezoidal pattern for determining envelope nonlinearity.

8.3 PRACTICAL ANALOG MODULATORS AND MULTIPLIERS

Although the need for analog multipliers seems to be universal, the existing circuits that are capable of accomplishing the task are few and far between. In addition, most practical multipliers suffer from either significant distortion in one or both input channels, which limits the dynamic range, or a poor frequency response in one or both channels. For example, a Hall multiplier, which is shown in Fig. 8.3–1, produces a transverse voltage V_t which is directly proportional to the longitudinal current

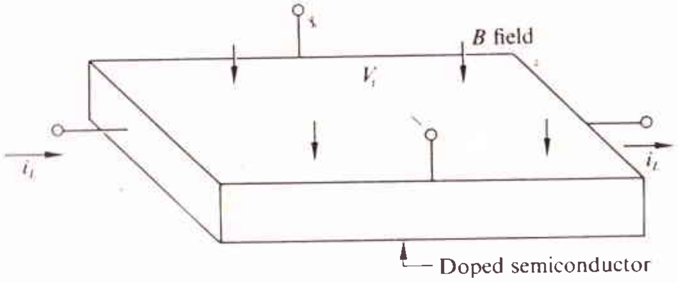


Fig. 8.3–1 Hall multiplier.

i_L and the magnitude of the normal magnetic induction field B_n, that is, $V_t = ki_LB_n$. Although perfect multiplication seems possible, the normal B field is severely limited in its maximum frequency. This limitation is due to the fact that the voltage amplitude required to produce the necessary B field increases linearly with frequency; thus unless special high-voltage circuitry is employed, B_n must be limited in frequency to under 10 kHz for normal Hall multipliers.

As a second example, an N-channel junction FET with a small drain-to-source voltage may be closely modeled as a voltage-controlled conductance of the form

$$g_{DS} = \frac{2I_{DSS}}{-V_P}\left(1 - \frac{v_{GS}}{V_P}\right), \qquad |v_{DS}| < 100\text{ mV},\ v_{GS} \geq V_P, \tag{8.3–1}$$

where g_{DS} is the drain-to-source conductance, I_{DSS} is the drain current with $v_{GS} = 0$ and $v_{DS} = -V_P$, and V_P is the pinch-off voltage. If this FET is placed in the input of

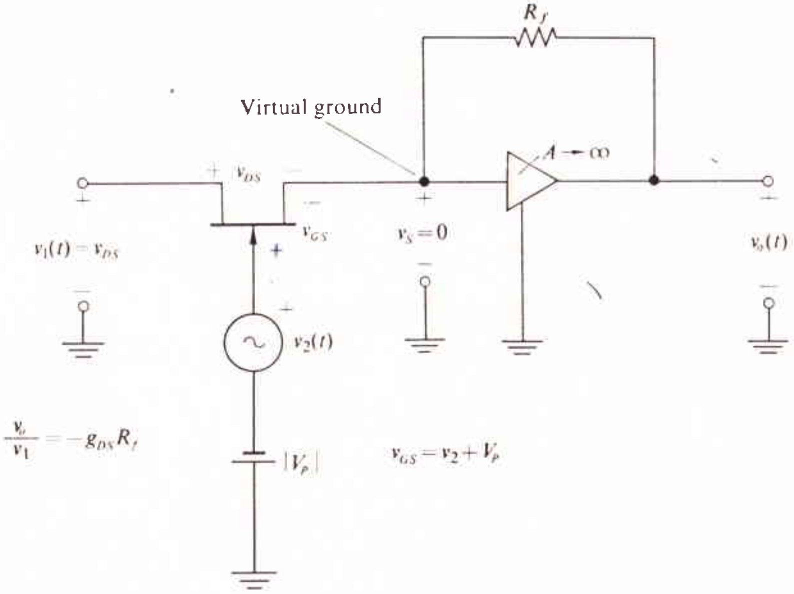


Fig. 8.3–2 FET multiplier.

a high-gain operational amplifier as shown in Fig. 8.3–2, then the output voltage $v_o(t)$ may be written as

$$v_o(t) = -g_{DS}R_fv_1(t)$$

$$= -R_f\frac{2I_{DSS}}{V_P}\frac{v_2(t)}{V_P}v_1(t), \tag{8.3–2}$$

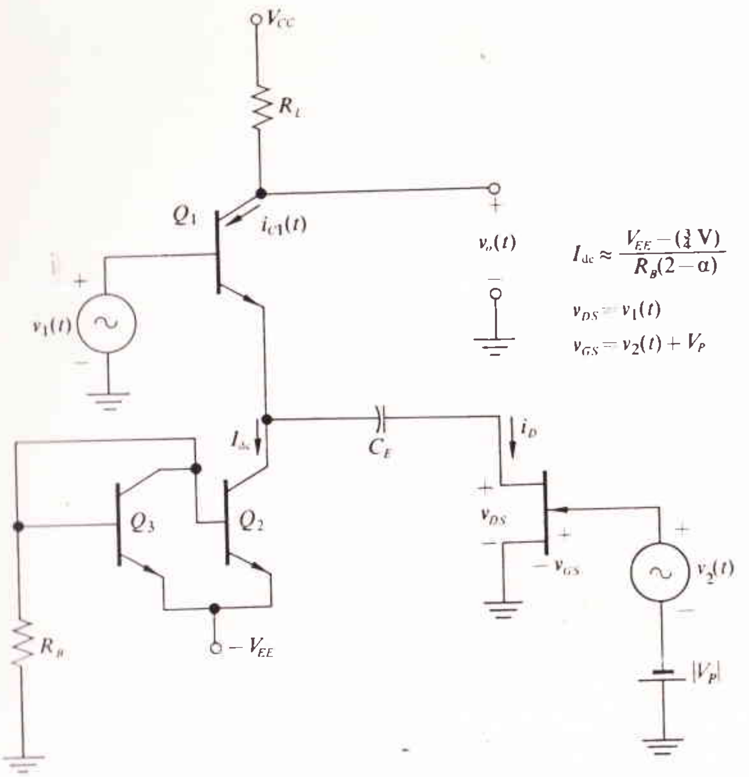


Fig. 8.3–3 FET-transistor multiplier.

which is proportional to the product of $v_1(t)$ and $v_2(t)$. However, this FET multiplier is restricted in dynamic range. The voltage $|v_1(t)| = |v_{DS}|$ must be kept less than 100 mV to limit distortion, and $v_2(t)$ is restricted to positive values less than $|V_P| +$ 0.7 V. Negative values of v_2 would cut off the FET, while positive values in excess of $|V_P| + 0.7$ V would turn on the gate-to-source diode.

An even simpler version of the multiplier of Fig. 8.3–2 may be obtained by replacing the operational amplifier by a single transistor, as shown in Fig. 8.3–3. In this circuit C_E is chosen sufficiently large to be a short circuit for the frequency components of $v_1(t)$, and I_{dc} is chosen sufficiently large so that $g_{in} = qI_{dc}/kT \gg g_{DS\max}$; hence $v_1(t)$ appears directly across the FET. In addition, no dc voltage appears across the FET because of C_E; thus

$$v_{DS} = v_1(t),$$

$$i_D(t) = g_{DS}v_1(t)$$

$$i_{C1}(t) = \alpha[I_{dc} + i_D(t)],$$

and finally,

$$\begin{aligned} v_o(t) &= V_{CC} - \alpha I_{dc}R_L - \alpha g_{DS}R_L v_1(t) \\ &= V_{CC} - \alpha I_{dc}R_L - \frac{2\alpha R_L I_{DSS}}{V_P} v_1(t)\frac{v_2(t)}{V_P}. \end{aligned} \tag{8.3–3}$$

It is apparent that $v_o(t)$ contains a term which is directly proportional to the product of $v_1(t)$ and $v_2(t)$: however, as in the case of the operational amplifier-FET multiplier, $|v_1(t)| = |v_{DS}|$ must be kept less than 100 mV while $v_2(t)$ is restricted to positive values less than $|V_P| + 0.7$ V.

Multipliers which employ mechanical shaft rotation of a linear potentiometer to produce multiplication, as shown in Fig. 8.3–4, are also limited by the maximum frequency at which the shaft can rotate and by the fact that x is constrained to be positive.

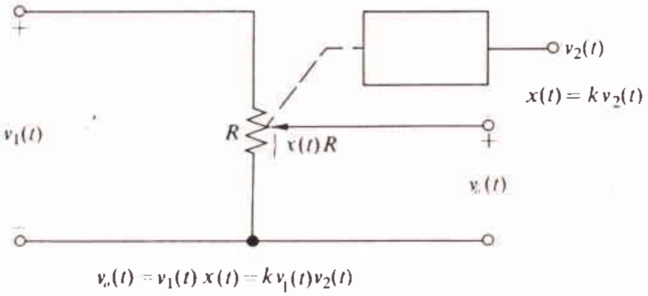


Fig. 8.3–4 Potentiometer multiplier.

The advent of integrated circuits, however, has now made possible an analog multiplier in which both inputs have a linear range which is a good fraction of the supply voltage and also have a frequency characteristic which extends into the gigahertz region. A typical integrated multiplier, introduced by Gilbert† and making use of the differential pair as its basic building block, is shown in skeleton form in Fig. 8.3–5. For this circuit, if $I_{ES1} = I_{ES2}$ and $I_{ES3} = I_{ES4}$ (a condition which requires the transistors to be integrated on the same chip and to have the same geometries), then from Eq. (4.6–3)

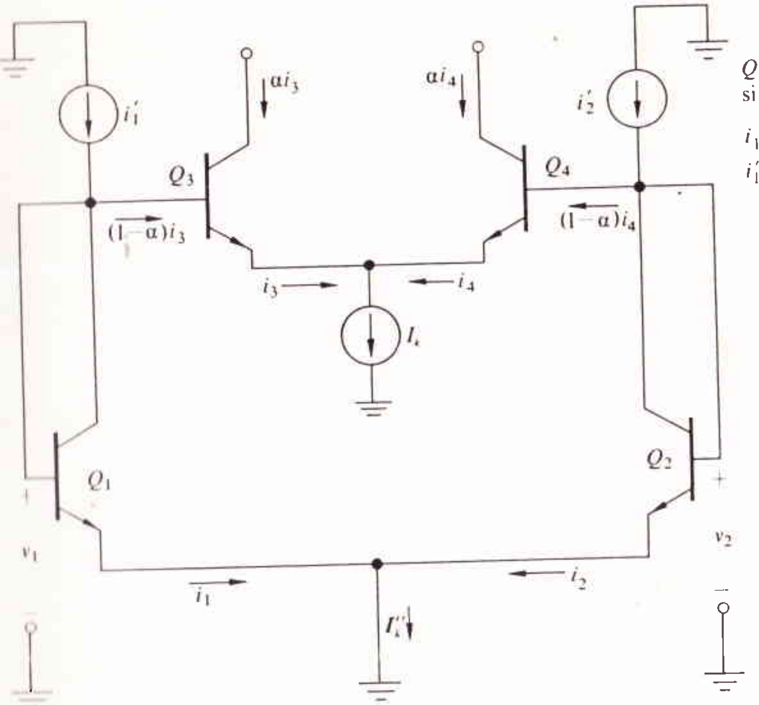
$$\frac{i_1}{I_k''} = \frac{i_3}{I_k} = \frac{1}{1+e^z} \qquad \text{and} \qquad \frac{i_2}{I_k''} = \frac{i_4}{I_k} = \frac{1}{1+e^{-z}}, \tag{8.3–4}$$

where $z = q(v_1 - v_2)kT$ and $I_k'' = i_1 + i_2$. In addition, since $i_1' = i_1 + (1-\alpha)i_3$,

$$i_1' = \frac{I_k'' + (1-\alpha)I_k}{1+e^z} = \frac{I_k'}{1+e^z}, \tag{8.3–5}$$

---

† B. Gilbert, "A Precise Four-Quadrant Multiplier with Subnanosecond Response." *IEEE Journal of Solid State Circuits*, SC-3, No. 4, pp. 365–373 (Dec. 1968).



Q_1 and Q_2 are active,
since $v_{CE} \approx V_0 > 0$

$i_1 + i_2 = I_k''$
$i_1' + i_2' = I_k'$
$I_k' = I_k'' + (1-\alpha)I_k$

Fig. 8.3–5 Differential-pair multiplier.

where $I_k' = I_k'' + (1 - \alpha)I_k = i_1' + i_2'$. Therefore, combining Eqs. (8.3–4) and (8.3–5), we observe that

$$i_3 = \frac{I_k i_1'}{I_k'} \tag{8.3–6}$$

and in a similar fashion

$$i_4 = \frac{I_k i_2'}{I_k'}. \tag{8.3–7}$$

It is apparent that, when the two differential pairs are combined as shown in Fig. 8.3–5, i_3 varies in proportion to i_1' while i_4 varies in proportion to i_2'. In addition, both i_3 and i_4 vary directly with I_k; thus if we constrain I_k' to be a constant (we can accomplish this by developing i_1' and i_2' from the collectors of another differential pair biased with a current source I_k'/α), then i_3 is proportional to the product of i_1' and I_k while i_4 is the proportional to the product of i_2' and I_k.

An alternative, but somewhat more practical, form of the differential-pair multiplier is shown in skeleton form in Fig. 8.3–6. As can readily be seen from the figure, i_1' and i_2' are outward-flowing current sources which, unlike the corresponding sources of Fig. 8.3–4, may be developed from an NPN differential pair. The ability

---

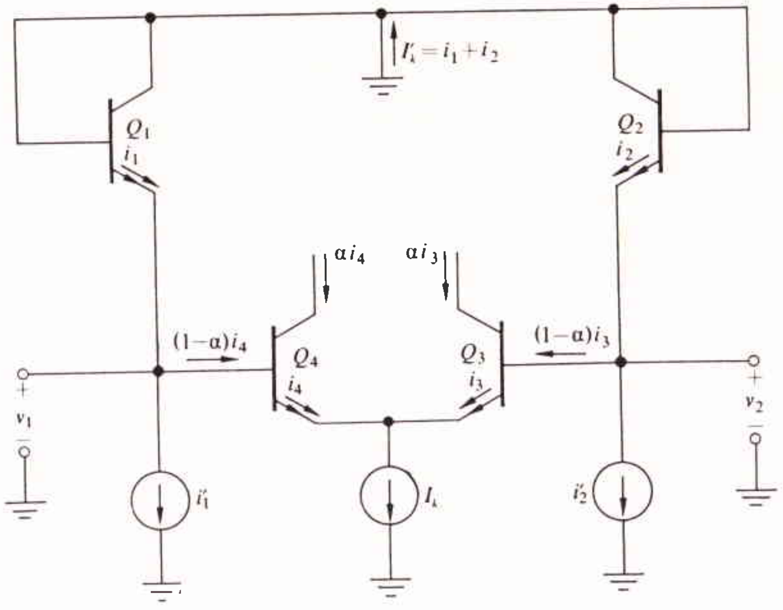


Fig. 8.3–6 Alternative form of differential-pair multiplier.

to employ all NPN transistors in a circuit greatly simplifies the mass integration process. The penalty which must be paid for this advantage is that the multiplier requires large values of $\beta = \alpha/(1 - \alpha)$, usually in excess of 150, for its proper operation.

For the differential-pair multiplier of Fig. 8.3–6 it is again readily demonstrated that if $I_{ES1} = I_{ES2}$ and $I_{ES3} = I_{ES4}$, then

$$\frac{i_1}{I_k'} = \frac{i_3}{I_k} = \frac{1}{1 + e^{z}} \qquad \frac{i_2}{I_k'} = \frac{i_4}{I_k} = \frac{1}{1 + e^{-z}},$$

where $z = q(v_1 - v_2)/kT$. In this configuration it is the currents in the diagonally opposite transistors which are proportional. Now if $(1 - \alpha)i_4 \ll i_1$ and $(1 - \alpha)i_3 \ll i_2$ (this is the high-β requirement), then

$$i_1 = i_1', i_2 = i_2', i_3 = \frac{I_k i_1'}{I_k'}, \qquad \text{and} \qquad i_4 = \frac{I_k i_2'}{I_k'}. \tag{8.3–8}$$

These equations are identical to Eqs. (8.3–6) and (8.3–7). Here again, if $I_k = i_1 + i_2 = i_1' + i_2'$ is constrained to be a constant, i_3 is proportional to the product of I_k and i_1'.

Since in both differential-pair multipliers i_1', i_2', and I_k must be positive (to prevent any transistor from being cut off†), the multipliers in their present form produce outputs lying within only one quadrant of a cartesian coordinate plane and are therefore referred to as a "one-quadrant multipliers." However either multiplier can be extended to a "two-quadrant multiplier" by taking advantage of the balanced nature of the differential pair.

---

† In addition, both i_1' and i_2' must be less than I_k' to prevent either Q_1 or Q_2 from being cut off.

If we define

$$i_{i1}(t) = 2i'_1 - I'_k = I'_k - 2i'_2, \tag{8.3–9}$$

or equivalently,

$$i'_1 = \frac{I'_k + i_{i1}}{2} \quad \text{and} \quad i'_2 = \frac{I'_k - i_{i1}}{2},$$

then $i_{i1}(t)$ has a dynamic range between $\pm I'_k$ and therefore occupies two quadrants. (This definition merely implies a dc shift in i'_i, which occurs naturally if a differential pair develops i'_1 and i'_2.) In terms of $i_{i1}(t)$, Eqs. (8.3–6) and (8.3–7) (or Eq. 8.3–8) reduce to

$$i_3 = \frac{I_k}{2}\left(\frac{i_{i1}}{I'_k} + 1\right) \quad \text{and} \quad i_4 = \frac{I_k}{2}\left(\frac{-i_{i1}}{I'_k} + 1\right). \tag{8.3–10}$$

By subtracting i_4 from i_3, which is readily accomplished at the collector of the differential pair, we obtain

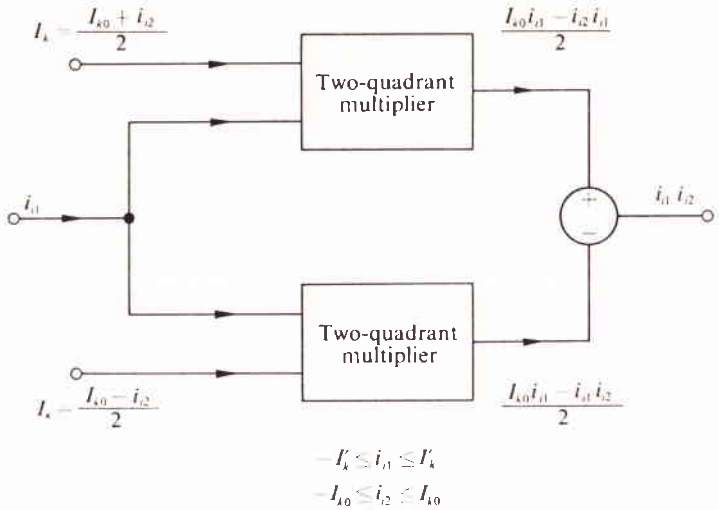
$$i_o = \alpha(i_3 - i_4) = \alpha\frac{I_k}{I'_k}i_{i1}, \tag{8.3–11}$$

which is a two-quadrant multiplication of $I_k(t)$ and $i_{i1}(t)$. A voltage $v_o = i_oR$ can be developed differentially by placing resistors of value R in the collectors of Q_3 and Q_4 in a fashion similar to that shown in Fig. 8.3–8 or using the scheme illustrated in Problem 8.13.

In general, two-quadrant multiplication is sufficient for generating normal AM, since $g(t)$ is always greater than zero and can be applied to the I_k input. However, for suppressed carrier AM, a "four-quadrant multiplier" is required.

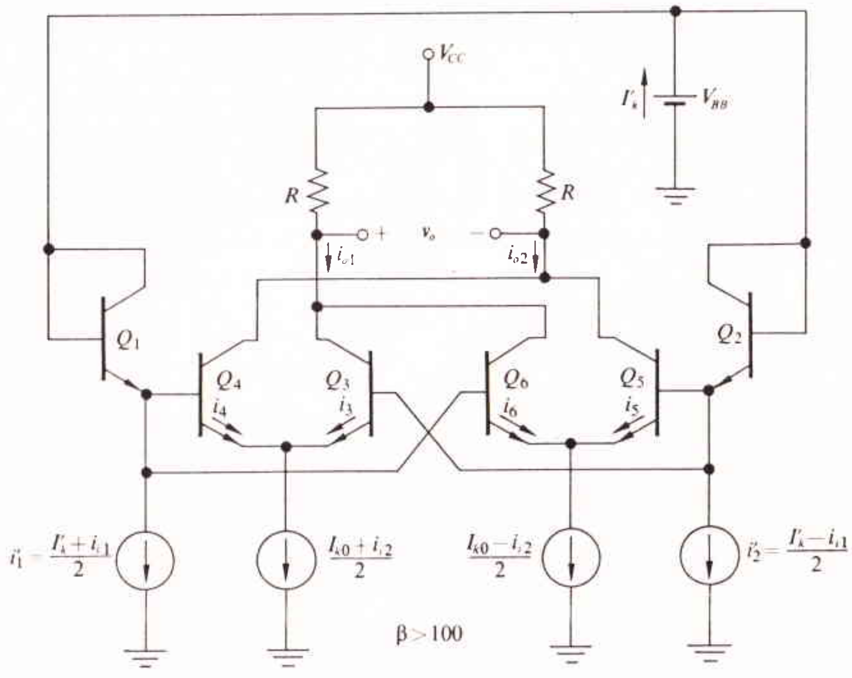
A four-quadrant multiplier can be constructed from two two-quadrant multipliers by connecting them as shown in Fig. 8.3–7. A corresponding four-quadrant differential-pair multiplier is shown in Fig. 8.3–8. For the multiplier of Fig. 8.3–8, both the Q_3–Q_4 pair and the Q_5–Q_6 pair share common base voltages with the Q_1–Q_2 pair; hence if β is high both i_3 and i_5 are proportional to i'_1 as well as their respective emitter bias suppliers, and i_4 and i_6 are proportional to i'_2 as well as their respective bias supplies. Therefore, from Eq. (8.3–8) we obtain

$$\begin{aligned}
i_3 &= \frac{I_{k0}}{4}\left(1 + \frac{i_{i2}}{I_{k0}}\right)\left(1 + \frac{i_{i1}}{I'_k}\right),\\
i_4 &= \frac{I_{k0}}{4}\left(1 + \frac{i_{i2}}{I_{k0}}\right)\left(1 - \frac{i_{i1}}{I'_k}\right),\\
i_5 &= \frac{I_{k0}}{4}\left(1 - \frac{i_{i2}}{I_{k0}}\right)\left(1 + \frac{i_{i1}}{I'_k}\right),\\
i_6 &= \frac{I_{k0}}{4}\left(1 - \frac{i_{i2}}{I_{k0}}\right)\left(1 - \frac{i_{i1}}{I'_k}\right),
\end{aligned} \tag{8.3–12}$$



$I_k = \frac{I_{k0} + i_{i2}}{2}$
$\frac{I_{k0}i_{i1} + i_{i2}i_{i1}}{2}$
Two-quadrant multiplier
i_{i1}
$i_{i1}i_{i2}$
+
−
Two-quadrant multiplier
$I_k = \frac{I_{k0} - i_{i2}}{2}$
$\frac{I_{k0}i_{i1} - i_{i1}i_{i2}}{2}$
$-I'_k \le i_{i1} \le I'_k$
$-I_{k0} \le i_{i2} \le I_{k0}$

Fig. 8.3–7 Synthesis of four-quadrant multiplier.



V_{CC}
I'_k
V_{BB}
R
R
+ v_o −
i_{o1}
i_{o2}
Q_1
Q_2
Q_4
i_4
Q_3
i_3
Q_6
i_6
Q_5
i_5
$i'_1 = \frac{I'_k + i_{i1}}{2}$
$\frac{I_{k0} + i_{i2}}{2}$
$\frac{I_{k0} - i_{i2}}{2}$
$i'_2 = \frac{I'_k - i_{i1}}{2}$
$\beta > 100$

Fig. 8.3–8 Four-quadrant differential-pair multiplier.

and in turn

$$i_{o1} = \alpha(i_3 + i_6) = \frac{\alpha I_{k0}}{2}\left(1 + \frac{i_{i1}i_{i2}}{I'_k I_{k0}}\right),$$

$$i_{o2} = \alpha(i_4 + i_5) = \frac{\alpha I_{k0}}{2}\left(1 - \frac{i_{i1}i_{i2}}{I'_k I_{k0}}\right). \tag{8.3–13}$$

Finally, provided that $i_{o1}R < V_{CC} - V_{BB} + 2V_0$ and $i_{o2}R < V_{CC} - V_{BB} + 2V_0$ to keep the differential pairs from saturating ($V_0 \approx \frac{3}{4}$ V for integrated silicon),

$$v_o = R(i_{o1} - i_{o2}) = \frac{\alpha R}{I'_k} i_{i1} i_{i2}, \tag{8.3–14}$$

where $-I'_k \leq i_{i1} \leq I'_k$ and $-I_{k0} \leq i_{i2} \leq I_{k0}$; thus we see that the circuit of Fig. 8.3–8 functions as a four-quadrant multiplier with a scale factor $K = \alpha R/I'_k$.

Since i_{i1} may be as large as I'_k and i_{i2} may be as large as I_{k0}, we observe from Eq. (8.3–13) that i_{o1} has a maximum value of αI_{k0}. Similarly i_{o2} has a maximum value of αI_{k0}. Therefore, a sufficient condition to ensure that the differential pairs do not saturate is that

$$\alpha I_{k0} R < V_{CC} - V_{BB} + 2V_0. \tag{8.3–15}$$

Figure 8.3–9 illustrates a typical circuit for developing the required differential currents to drive the multiplier of Fig. 8.3–8.† In this circuit, if $R_E \gg r_{in} = 2kT/qI_k$, which is usually the case, then

$$i_{C1} = \alpha i_{E1} = \alpha\left[\frac{I_{dc}}{2} + \frac{v_i(t)}{R_E}\right] \quad \text{and} \quad i_{C2} = \alpha i_{E2} = \alpha\left[\frac{I_{dc}}{2} - \frac{v_i(t)}{R_E}\right] \tag{8.3–16}$$

for

$$\frac{I_{dc}}{2} \leq \frac{v_i(t)}{R_E} \leq \frac{I_{dc}}{2}.$$

In addition, if Q_3, Q_4, and Q_5 are identical in geometry, then $I_{E3} = I_{E4} = I_{E5} = I_E$ and hence a net current of $(3 - 2\alpha)I_E$ flows through R_B. Thus

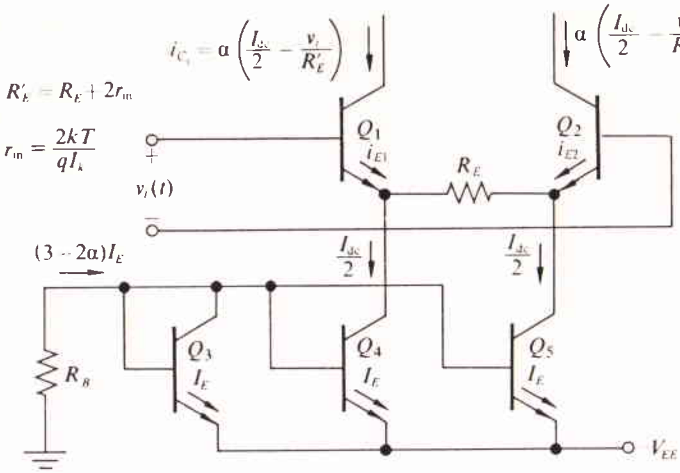
$$\frac{I_{dc}}{2} = \alpha I_E = \frac{\alpha(V_{EE} - V_0)}{R_B(3 - 2\alpha)} \approx \frac{V_{EE} - V_0}{R_B}. \tag{8.3–17}$$

Figure 8.3–10 illustrates a complete integrated multiplier which includes the differential current source drivers. In this circuit (if $R_E \gg r_{in}$) we can identify

$$i_{i1} = \frac{2\alpha v_{i1}}{R_{E1}}, \qquad i_{i2} = \frac{2\alpha v_{i2}}{R_{E2}}, \qquad I_{k0} = \alpha I_{dc2}, \qquad \text{and} \qquad I'_k = \alpha I_{dc1};$$

---

† An alternative circuit form would have two series emitter resistors of value $R_E/2$ with a single current source driving their common node. This approach suffers from two disadvantages in comparison with the circuit of Fig. 8.3–8. First, an additional resistor is needed which requires more area on an integrated chip than the transistor it replaces; second, the bias current $I_{dc}/2$ flows through the resistor, making the Q-point a function of R_E.



$$I_{ES3} = I_{ES4} = I_{ES5} \rightarrow I_{E3} = I_{E4} = I_{E5} = I_E$$

$$\frac{I_{dc}}{2} = \frac{\alpha(V_{EE} - V_0)}{(3 - 2\alpha)R_B} \approx \frac{V_{EE} - V_0}{R_B}, \ \beta \gg 1$$

Fig. 8.3–9 Current drive for differential-pair multiplier.

hence, provided that all transistors remain in their active region, from Eq. (8.3–14) the four-quadrant multiplier output becomes

$$v_o(t) = \frac{4\alpha^2 R}{R_{E1}R_{E2}I_{dc1}} v_{i1}(t)v_{i2}(t), \tag{8.3–18}$$

which is the desired product of the two input voltages $v_{i1}(t)$ and $v_{i2}(t)$.

To ensure the validity of Eq. (8.3–18), the following restrictions on the magnitudes of $v_{i1}(t)$, $v_{i2}(t)$, and I_{dc2} must be observed:

1) $-V_{EE} + V_0 < v_{i1} < V_{BB}$ to keep Q_7 active,
2) $-V_{EE} + V_0 < v_{i2} < V_{BB} - V_0$ to keep Q_9 active,
3) $\frac{-I_{dc1}}{2} \leq \frac{v_{i1}}{R_{E1}} \leq \frac{I_{dc1}}{2}$ to keep Q_7 and Q_8 from being cut off,
4) $\frac{-I_{dc2}}{2} \leq \frac{v_{i2}}{R_{E2}} \leq \frac{I_{dc2}}{2}$ to keep Q_9 and Q_{10} from being cut off,
5) $\alpha^2 I_{dc2} R < V_{CC} - V_{BB} + 2V_0$ to keep Q_3, Q_4, Q_5, or Q_6 from saturating (cf. Eq. 8.3–15).

In most integrated analog multipliers of the form shown in Fig. 8.3–10, the transistors are integrated on one chip; however, the resistors R, R_{B1}, R_{B2}, R_{E1}, R_{E2}, and R_{BB} must be supplied externally by the engineer making use of the multiplier.

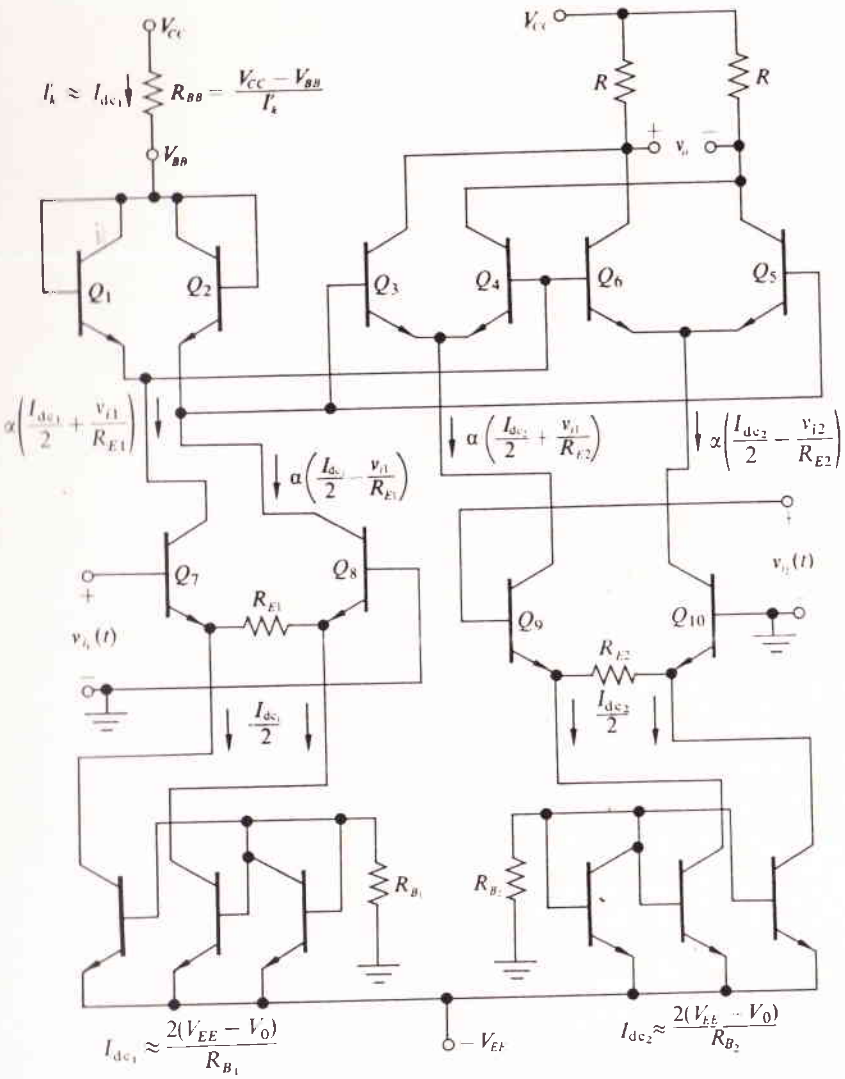

V_{CC}
$I_k \approx I_{dc_1}$
$R_{BB} = \frac{V_{CC} - V_{BB}}{I_k}$
V_{BB}
V_{CC}
R
R
+ v_o −
Q_1
Q_2
Q_3
Q_4
Q_6
Q_5
$\alpha\left(\frac{I_{dc_1}}{2} + \frac{v_{i1}}{R_{E1}}\right)$
$\alpha\left(\frac{I_{dc_2}}{2} + \frac{v_{i1}}{R_{E2}}\right)$
$\alpha\left(\frac{I_{dc_2}}{2} - \frac{v_{i2}}{R_{E2}}\right)$
$\alpha\left(\frac{I_{dc_1}}{2} - \frac{v_{i1}}{R_{E1}}\right)$
Q_7
Q_8
R_{E1}
+
$v_{i1}(t)$
$v_{i2}(t)$
Q_9
Q_{10}
R_{E2}
$\frac{I_{dc_1}}{2}$
$\frac{I_{dc_2}}{2}$
R_{B_1}
R_{B_2}
$I_{dc_1} \approx \frac{2(V_{EE} - V_0)}{R_{B_1}}$
$-V_{EE}$
$I_{dc_2} \approx \frac{2(V_{EE} - V_0)}{R_{B_2}}$

Fig. 8.3–10 Complete integrated multiplier.

In general, R is chosen sufficiently small so that it is not appreciably loaded by either the input impedance of the next differential pair or stray output capacitance. For example, if 20 pF of output capacitance exists and the multiplier is to operate up to frequencies of 20 MHz, then

$$R < \frac{1}{2\pi(20 \times 10^6)(20 \times 10^{-12})} = 400\ \Omega.$$

On the other hand, R is not chosen any smaller than necessary; otherwise, enormous currents would be required to develop a reasonable output voltage.

On the assumption that V_{CC} is specified at 10 V, $-V_{EE}$ is specified at -5 V, and

$$|v_{i1}|_{\max} = |v_{i2}|_{\max} = 4\ \text{V},$$

V_{BB} is chosen to be greater than 4 V + 0.75 V = 4.75 V to avoid saturating Q_9. A choice of 5 V in this case would be reasonable. With $R = 400\ \Omega$ and $V_{BB} = 5$ V, restriction (5) limits I_{dc2} to

$$\frac{1}{\alpha^2}\frac{6.5\ \text{V}}{400\ \Omega} \approx 16.3\ \text{mA}.$$

Probably a value of 16 mA would be selected, since a smaller value would require a larger value of R_{E2} [restriction (4)] and thus a smaller output voltage.

With $I_{dc2} = 16$ mA the minimum value of R_{E2} is

$$R_{E2} = \frac{2|v_{i2}|_{\max}}{I_{dc2}} = 500\ \Omega$$

(a standard value of 499Ω would be chosen). Now for maximum output voltage we also desire $I_{dc1}R_{E1}$ to be as small as possible; however, $I_{dc1}R_{E1}$ is constrained to be greater than $2|v_{i1}|_{\max} \approx 8$ V. A choice of $I_{dc1} = 9$ mA and $R_{E1} = 1\ \text{k}\Omega$ is reasonable. With these values we can be sure that all transistors operate in their active region and that

$$v_o(t) = \frac{1}{2.87\ \text{V}} v_{i1}(t) v_{i2}(t). \tag{8.3–19}$$

Finally, we choose

$$R_{BB} = \frac{5\ \text{V}}{9\ \text{mA}} \approx 5.6\ \text{k}\Omega$$

to yield the desired value for V_{BB},

$$R_{B1} = \frac{4.25\ \text{V}}{9\ \text{mA}} \approx 470\ \Omega$$

to yield the desired value of I_{dc1}, and

$$R_{B2} = \frac{4.25\ \text{V}}{16\ \text{mA}} \approx 270\ \Omega$$

to yield the desired value of I_{dc2}.

Now, for this multiplier, if

$$v_{i1}(t) = (4\ \text{V}) f(t) \qquad \text{and} \qquad v_{i2}(t) = (4\ \text{V}) \cos 10^8 t,$$

then the output $v_o(t)$ takes the form

$$v_o(t) = (5.57\ \text{V}) f(t) \cos 10^8 t,$$

which is a suppressed carrier AM signal. Clearly the scale factor could be decreased to any value less than 5.57 V by increasing either R_{E1} or R_{E2} or by decreasing R.

nalog Multipliers with One Nonlinear Input Channel

Many multipliers exist in which one input channel is linear while the other input hannel is highly nonlinear. An example of this type of multiplier is a single differential air for which the collector current is directly proportional to I_k and a highly onlinear function of $v_1 - v_2$ (Eq. 4.6–3). Although multiplication in such a device s not ideal, amplitude modulation can be accomplished by applying the carrier to he nonlinear input channel and the modulation to the linear input channel while placing the output through a bandpass filter centered about the carrier. The bandpass filter removes the harmonics of the carrier generated by the nonlinear channel.

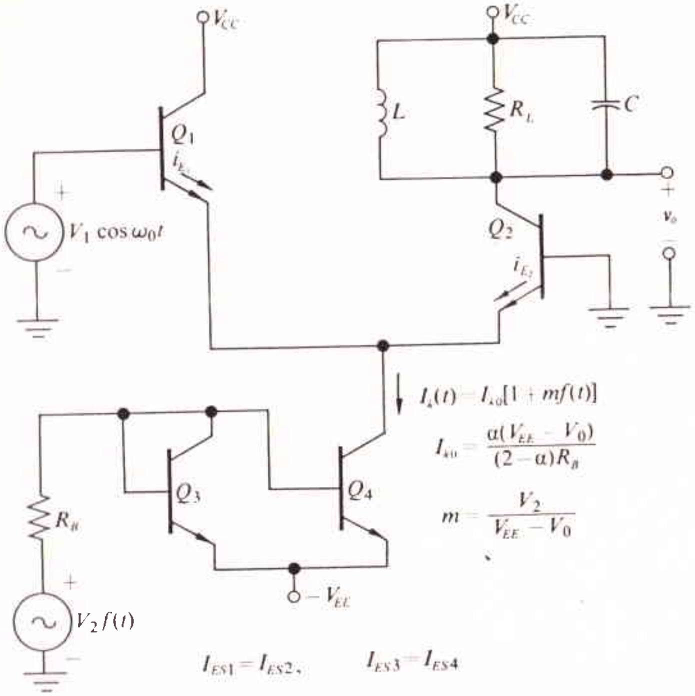

V_{CC}
V_{CC}
L
R_L
C
Q_1
i_{E_1}
+
$V_1 \cos \omega_0 t$
−
Q_2
i_{E_2}
+
v_o
$I_k(t) = I_{k0}[1 + mf(t)]$
$I_{k0} = \frac{\alpha(V_{EE} - V_0)}{(2-\alpha)R_B}$
$m = \frac{V_2}{V_{EE} - V_0}$
Q_3
Q_4
R_B
+
$V_2 f(t)$
−
$-V_{EE}$
$I_{ES1} = I_{ES2}$, $I_{ES3} = I_{ES4}$

Fig. 8.3–11 Differential-pair AM modulator.

Such a modulator employing an integrated differential pair is shown in Fig. 8.3–11. For this modulator, Q_3 and Q_4 act as a current source to develop a current drive $I_k(t)$ to the differential pair of the form

$$I_k(t) = I_{k0}[1 + mf(t)], \tag{8.3–20}$$

where $I_{k0} = \alpha(V_{EE} - V_0)/(2 - \alpha)R_B$ and $m = V_2/(V_{EE} - V_0)$. This assumes, of course, that Q_3 and Q_4 are integrated on the same chip and with identical geometries, so that $I_{ES3} = I_{ES4}$.) In addition, from Eqs. (4.6–9) and (4.6–10),

$$i_{E2} = I_k(t)[\tfrac{1}{2} - a_1(x) \cos \omega_0 t - a_3(x) \cos \omega_0 t - \cdots], \tag{8.3–21}$$

where $x = qV_1/kT$ and $a_1(x)$ is the Fourier coefficient plotted in Fig. 4.6–4. Therefore, the fundamental component of the collector current of Q_2 is given by

$$i_{C21} = -I_{k0}a_1(x)[1 + mf(t)] \cos \omega_0 t, \tag{8.3–22}$$

which results in

$$\begin{aligned} v_o(t) &= V_{CC} + I_{k0}a_1(x)\{[1 + mf(t)] \cos \omega_0 t\} * z_{11}(t) \\ &= V_{CC} + I_{k0}a_1(x)\{[1 + mf(t)] * z_{11L}(t)\} \cos \omega_0 t, \end{aligned} \tag{8.3–23}$$

where $z_{11}(t)$ is the impulse response of the output tuned circuit and $z_{11L}(t)$ is the low-pass equivalent of $z_{11}(t)$ (cf. Fig. 3.1–3). Implicit in the expression for $v_o(t)$ given by Eq. (8.3–22) is the fact that the tuned circuit has a sufficiently high Q to remove the low-frequency and higher-harmonic components of $i_{C2}(t)$. If, in addition, the output filter is flat over the band of frequencies occupied by $i_{C21}(t)$, then $v_o(t)$ simplifies to

$$v_o(t) = V_{CC} + I_{k0}R_La_1(x)[1 + mf(t)] \cos \omega_0 t, \tag{8.3–24}$$

which is clearly the form of the desired AM signal.

As a numerical example let

$$\begin{array}{llll} V_{CC} = 12\text{ V}, & V_{EE} = -12\text{ V}, & V_1 = 1.56\text{ mV}, & V_2 = 5.63\text{ V}, \\ R_B = 6\text{ k}\Omega, & R_L = 5\text{ k}\Omega, & C = 1000\text{ pF}, & L = 10\ \mu\text{H}, \\ \omega_0 = 10^7\text{ rad/sec}, & \text{and} & f(t) = \cos 10^5 t, & \end{array}$$

from which it follows that

$$m = \tfrac{1}{2}, \qquad I_{k0} = \frac{11.25\text{ V}}{6\text{ k}\Omega} = 1.875\text{ mA}, \qquad x = 6,$$

and from Fig. 4.6–4,

$$a_1(6) = 0.6.$$

Therefore

$$i_{C21}(t) = (1.25\text{ mA})(1 + \tfrac{1}{2} \cos 10^5 t) \cos 10^7 t$$

and

$$\begin{aligned} v_o(t) &= (12\text{ V}) + (1.125\text{ mA})[(1 + \tfrac{1}{2} \cos 10^5 t) \quad z_{11L}(t) \cos 10^7 t] \\ &= (12\text{ V}) + (5.63\text{ V})\left[1 + \frac{1}{2\sqrt{2}} \cos\left(10^5 t - \frac{\pi}{4}\right)\right] \cos 10^7 t. \end{aligned}$$

The attenuation of the $\cos 10^5 t$ term by $1/\sqrt{2}$ and its phase shift by $-\pi/4$ results directly from the fact that the low-pass equivalent of the output-tuned circuit has a −3 dB bandwidth of 10^5 rad/sec.

Note that the minimum value of $v_o(t)$ is 4.4 V; hence Q_2 does not saturate.

PRACTICAL CHOPPER MODULATORS

e key component in a chopper modulator is the voltage-controlled switch which ens and closes at the carrier rate. Therefore, in this section we shall look at two ltage-controlled single-pole single-throw (SPST) switches—one employing a ode bridge and the other employing a single FET—and then we shall consider the oblem of employing these single-pole switches to construct a voltage-controlled versing switch. The SPST switch is used in the single-ended chopper modulator own in Fig. 8.2–2, whereas the reversing switch is employed in the balanced chopper odulator shown in Fig. 8.2–4.

Note that, although the voltage-controlled switch is being discussed in con-nction with chopper modulators, it functions equally well as a synchronous de-odulator or a mixer in the same configuration as Fig. 8.2–2. For a synchronous tector, however, the output filter must be replaced by a low-pass filter, whereas r a mixer the output bandpass filter must be tuned to the intermediate frequency.

iode-Bridge Modulator

lmost all single-ended chopper modulators in use today employ the diode bridge the voltage-controlled switch. Figure 8.4–1 illustrates a typical diode-bridge odulator in which a positive value of $v_1(t)$ causes all the bridge diodes to conduct ereby bringing $v_a(t)$ close to ground potential, and in which a negative value of v_1 everse-biases all the bridge diodes, thereby permitting $v_a(t)$ to follow $g(t)$.

It is apparent that the amplitude V_1 of $v_1(t)$ must be sufficiently large when v_1 negative to keep all the diodes reverse-biased. A little thought indicates that for $(t) \geq 0$, and with $v_1(t) = -V_1$, D_2 and D_3 are on the verge of conduction when $(t) = V_1 + 2V_0$; and that for $g(t) < 0$, D_1 and D_4 are on the verge of conduction for $(t)| = V_1 + 2V_0$. Hence to ensure that all the bridge diodes remain reverse-biased or $v_1 = -V_1$, we require that

$$V_1 > g(t) - 2V_0 \tag{8.4–1}$$

or all t.

It is also apparent that V_1 must be sufficiently large when $v_1 = +V_1$ to keep all he diodes forward-biased so that the bridge presents a low impedance to ground; hat is, $v_a(t)$ should be a small voltage with $v_1 = V_1$. To determine the required magni-ude of V_1 in this case, we define the current leaving the v_1 source as I_2 and the current eaving the $g(t)$ source as $i_1(t)$. Since, in general, the four bridge diodes are integrated on a single chip with identical geometries, the bridge is balanced and $i_1(t)$ and I_2 split equally between the two bridge arms; thus

$$i_{D1}(t) = \frac{I_2 - i_1(t)}{2} = i_{D4}(t),$$

$$i_{D2}(t) = \frac{I_2 + i_1(t)}{2} = i_{D3}(t). \tag{8.4–2}$$

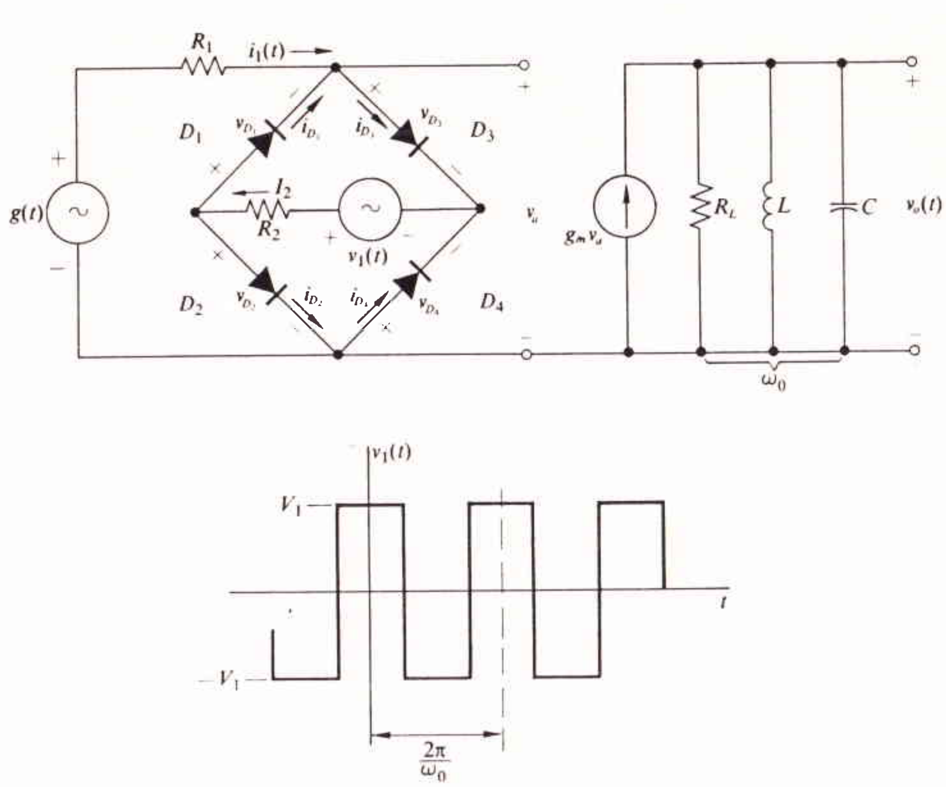


Fig. 8.4–1 Chopper modulator employing diode bridge.

If we assume that each diode is characterized by the volt-ampere relationship

$$i_D = I_S(e^{qv_D/kT} - 1) \approx I_S e^{qv_D/kT}, \tag{8.4–3}$$

then

$$\begin{aligned} v_a = v_{D2} - v_{D1} &= \frac{kT}{q}\left(\ln\frac{I_2 + i_1(t)}{2I_S} - \ln\frac{I_2 - i_1(t)}{2I_S}\right) \\ &= \frac{kT}{q}\ln\left(\frac{1 + i_1(t)/I_2}{1 - i_1(t)/I_2}\right). \end{aligned} \tag{8.4–4}$$

By expanding $v_a(t)$ in a MacLauren series in i_1/I_2, we obtain

$$v_a(t) = r_d i_1\left(1 + \frac{i_1^2}{3I_2^2} + \frac{i_1^4}{5I_2^4} + \cdots\right), \tag{8.4–5}$$

where $r_d = (kT/q)(2/I_2)$ is the small-signal diode resistance with $I_2/2$ as a bias current. It is apparent that, if we wish to keep nonlinear components of i_1 [which is proportional to the modulation $g(t)$] out of the output to avoid envelope distortion in $v_o(t)$, then $i_1^2/3I_2^2 \ll 1$. With this restriction $v_a(t) = r_d i_1(t)$ and the forward-biased diode bridge may be modeled as a single resistor of value r_d shunting v_a; hence the diode bridge takes the form of an ideal voltage-controlled switch in series with a resistance r_d.

With this model

$$i_1(t) = \frac{g(t)}{R_1 + r_d}$$

while

$$I_2 = \frac{V_1 - 2V_0}{R_2} \qquad \text{and} \qquad r_d = \frac{2R_2}{q(V_1 - 2V_0)/kT};$$

thus to keep $i_1^2/3I_2^2 < 0.01$,

$$V_1 > \frac{5.8g(t)R_2}{R_1 + r_d} + 2V_0 \tag{8.4–6}$$

for all t. The inequalities of Eqs. (8.4–1) and (8.4–6) may be satisfied simultaneously by choosing $V_1 > g(t) - 2V_0$ and choosing R_2 of the order of $\frac{1}{10}R_1$. For example, if $|g(t)|_{\max} = 10$ V, then the bridge remains open with $v_1 = -V_1$ for $V_1 > 8.5$ V ($V_0 = \frac{3}{4}$ V). If we select $V_1 = 9$ V, then if

$$R_2 = \frac{R_1 + r_d}{7.75} \approx \frac{R_1}{7.75},$$

the bridge appears as a resistor r_d with $v_1 = V_1$.

A complete diode-bridge modulator which incorporates the floating source $v_1(t)$ as well as the output filter is shown in Fig. 8.4–2. In this circuit the transformer

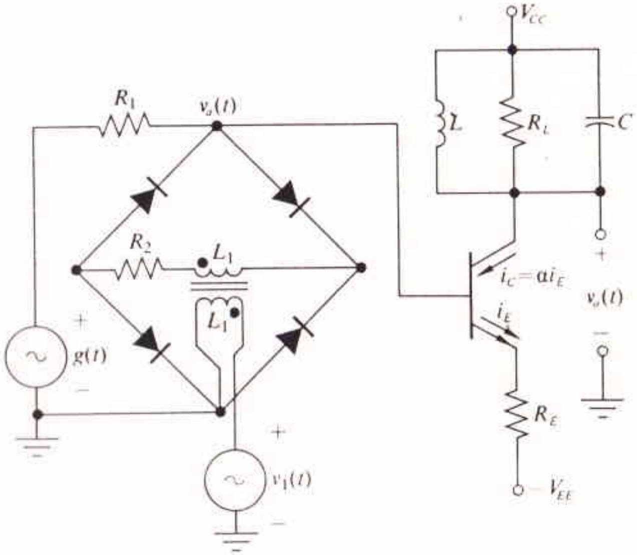


Fig. 8.4–2 Practical balanced modulator.

is a closely coupled transformer operating in its midband range, and therefore functions as an ideal transformer. In addition, $(1 + \beta)R_E$ is large in comparison with R_1 so that the transistor does not load the bridge. Consequently, if the output-tuned circuit is broad enough to pass the modulation and yet narrow enough to remove the low-frequency and higher-harmonic components of $v_a(t)$, then from Eq. (8.2–19)

$$v_o(t) = \frac{2g(t)}{\pi R_E} \frac{R_1}{R_1 + r_d} \cos \omega_0 t + V_{CC}. \tag{8.4–7}$$

Equation (8.4–7) assumes, of course, that Eqs. (8.4–1) and (8.4–6) have been satisfied and that the transistor remains in its active region.

The control voltage $v_1(t)$ may be supplied by a sufficiently large sine wave of radian frequency ω_0 instead of a square wave. If, as shown in Fig. 8.4–3, V_1 is large in comparison with V_α and V_β (Eqs. 8.4–1 and 8.4–6), then the sine wave functions in essentially the same fashion as the square wave in controlling the states of the bridge.

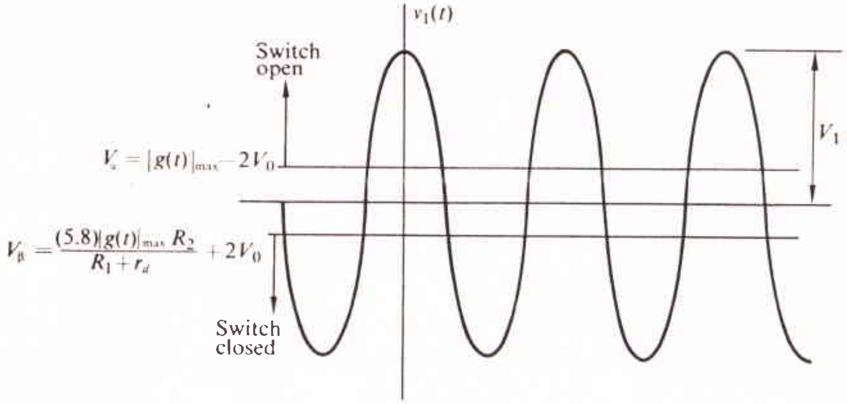

$v_1(t)$
Switch open
$V_\alpha = |g(t)|_{\max} - 2V_0$
V_1
$V_\beta = \frac{(5.8)|g(t)|_{\max} R_2}{R_1 + r_d} + 2V_0$
Switch closed

Fig. 8.4–3 Sinusoidal control voltage.

One main advantage of a sine-wave drive is that the transformer coupling $v_1(t)$ to the diode bridge need not be nearly as broadband. On the other hand, the larger value of V_1 with a sine-wave drive requires a much higher breakdown voltage for the bridge diodes.

Whether $v_1(t)$ is a sine wave or a square wave, in practical diode bridges short-duration transient "spikes" appear on $v_a(t)$ in the vicinity of the bridge transitions from open to closed because of parasitic capacitance and diode charge storage. These spikes are, in general, of little consequence, since they contain sufficiently high-frequency components so that they are not transmitted to the output through the bandpass filter $H(j\omega)$.

FET Modulator

A junction or an insulated-gate FET may be employed instead of the diode bridge as the voltage-controlled switch in a chopper modulator. Figure 8.4–4 illustrates a

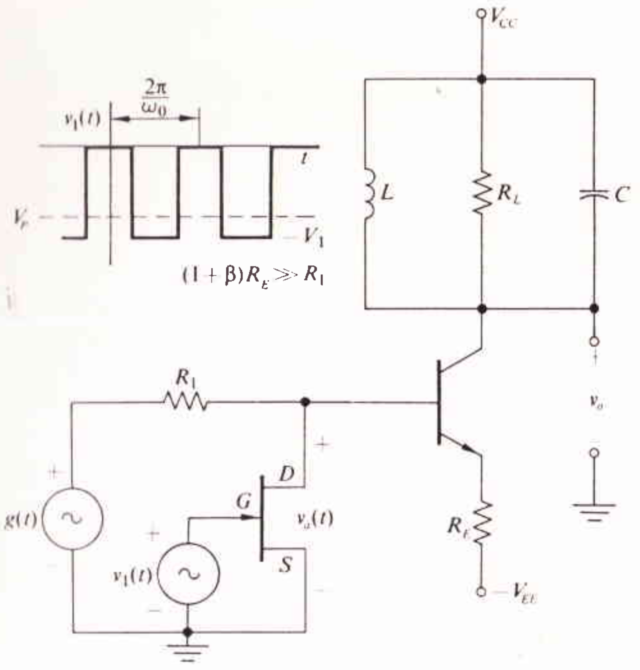


Fig. 8.4–4 N-channel junction FET chopper modulator.

ical N-channel junction FET chopper modulator. In this circuit, with $v_1(t) =$ $_1 < V_P$ (V_P is the pinch-off voltage of the FET), the FET opens, permitting $v_a(t)$ ollow $g(t)$. On the other hand, with $v_1(t) = 0$, then the FET functions as an ohmic ductance g_{DSS} of value

$$g_{DSS} = \frac{2I_{DSS}}{-V_P}\left(1 - \frac{v_{GS}}{V_P}\right)_{v_{GS}=0} = \frac{2I_{DSS}}{-V_P}, \tag{8.4–8}$$

vided that $|v_{DS}| = |v_a| < 100\,\text{mV}$. Consequently the FET may be modeled as an l voltage-controlled switch in series with a resistance $r_{DSS} = 1/g_{DSS}$. For typical tion and insulated-gate FET's, r_{DSS} varies from several ohms to several thousand ns.

With $v_1(t) = 0$,

$$v_a(t) = v_{DS} = \frac{g(t)r_{DSS}}{R_1 + r_{DSS}};$$

s to ensure that $|v_{DS}(t)|$ remains less than 100 mV for all t we require R_1 to be iciently large so that

$$R_1 > r_{DSS}\left(\frac{|g(t)|_{\max}}{100\,\text{mV}} - 1\right). \tag{8.4–9}$$

example, if $|g(t)|_{\max} = 5\,\text{V}$ and $r_{DSS} = 500\,\Omega$, then $R_1 > 24.5\,\text{k}\Omega$. To avoid ling by the output transistor when the FET is reverse biased, the resistor R_1

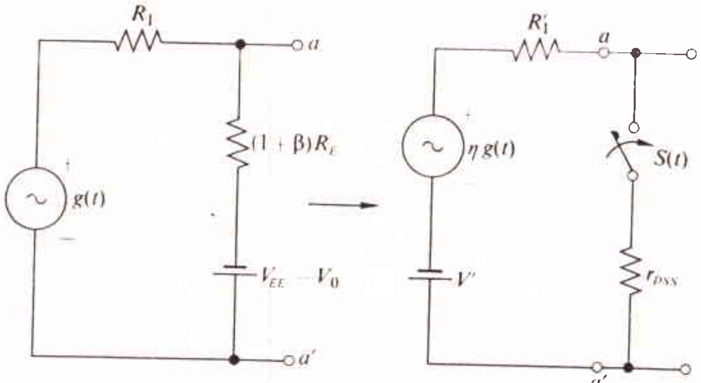
---

should not be chosen too much greater than this value. If, on the other hand, R_1 is chosen to be less than 24.5 kΩ, then $v_a(t)$ exceeds 100 mV, r_{DSS} becomes nonlinear, and $v_a(t)$ is no longer a linear function of $g(t)$; consequently nonlinear envelope distortion begins to appear on the output AM wave.

If R_E is sufficiently large so that transistor loading can be neglected, then $v_o(t)$ is given by (cf. Eq. 8.4–7)

$$v_o(t) = \frac{2g(t)}{\pi R_E}\,\frac{R_1}{R_1 + r_{DSS}}\cos\omega_0(t) + V_{CC}. \tag{8.4–10}$$

However, if R_E is not sufficiently large, then the loading must be incorporated with the $g(t)$-R_1 network as shown in Fig. 8.4–5 by forming a Thévenin equivalent network. Clearly $g(t)$ is decreased by a factor of η because of the loading; however, in addition, a dc bias V' is added in series with $g(t)$. If $\overline{g(t)} = 0$, as it is for suppressed carrier modulation, then the presence of V' produces an average component in the modulation voltage being chopped and thus a nonzero carrier at the output. To eliminate this undesired carrier component, either R_E must be increased relative to R_1 or an isolation stage such as a source follower must be inserted between the chopper stage and the output transistor.



$$\eta = \frac{(1+\beta)R_E}{(1+\beta)R_E + R_1}, \quad V' = \frac{(V_{EE} - V_0)R_1}{(1+\beta)R_E + R_1}, \quad R_1' = R_1 \ (1+\beta)R_E$$

Fig. 8.4–5 Effect of transistor loading.

In addition to the diode bridge or the FET chopper, a bipolar transistor being switched between saturation and cutoff may be employed as the voltage-controlled switch. However, when saturated, the transistor may be modeled as a resistor in series with a dc voltage source of approximately 100 mV (for silicon). This voltage source has the effect of introducing a carrier component at the modulator output, which is quite undesirable if suppressed carrier AM is being generated. This saturation voltage may be largely balanced out by placing two transistors in series (emitter to emitter) and placing the switching voltage between their bases.

anced Chopper Modulator

ure 8.4-6 indicates how two diode bridges can be employed to alternately apply (t) and $-g(t)$ across R_1 and thus produce the effect of a reversing switch. Note t the bridges are arranged so that one bridge is open when the other is closed. It pparent that the closed bridge is unaffected by the open bridge and thus Eq. -6) still determines the value of V_1 required to ensure that the bridge remains

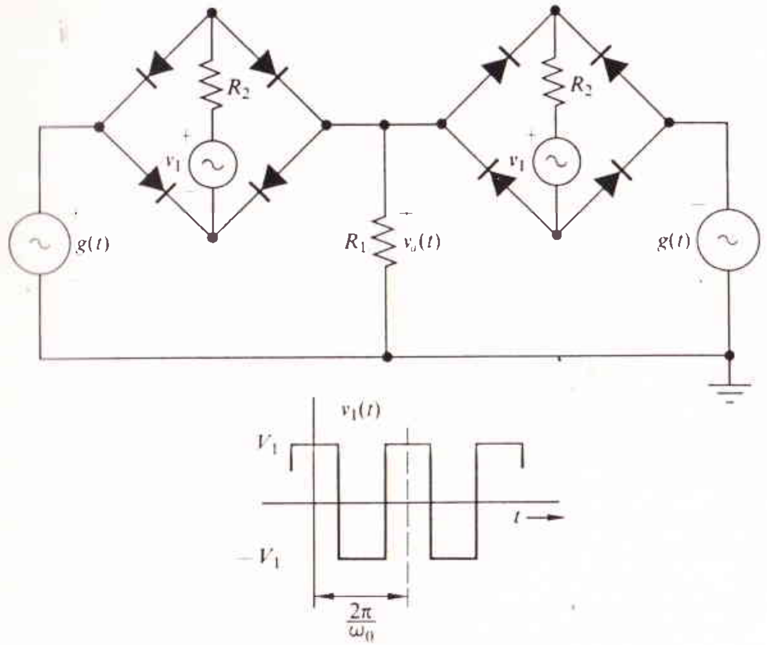


Fig. 8.4-6 Reversing switch for balanced chopper modulator.

sed for all t. On the other hand, the closed bridge does affect the open bridge in t it increases the voltage across the open bridge to

$$g(t)\left(1 + \frac{R_1}{R_1 + r_d}\right);$$

ice with the arguments employed to obtain Eq. (8.4–1) we require

$$V_1 > g(t)\left(1 + \frac{R_1}{R_1 + r_d}\right) - 2V_0 \qquad (8.4\text{–}11)$$

all t to ensure that all the diodes in the open bridge remain reverse biased.

Figure 8.4-7 illustrates a practical chopper modulator in which both $g(t)$ and) are supplied from grounded sources. The transformer T_1 is a closely coupled ter-tapped audio transformer with a midband frequency range sufficient to pass frequency components of $g(t)$ (see Section 2.2), while the transformer T_2 uses a

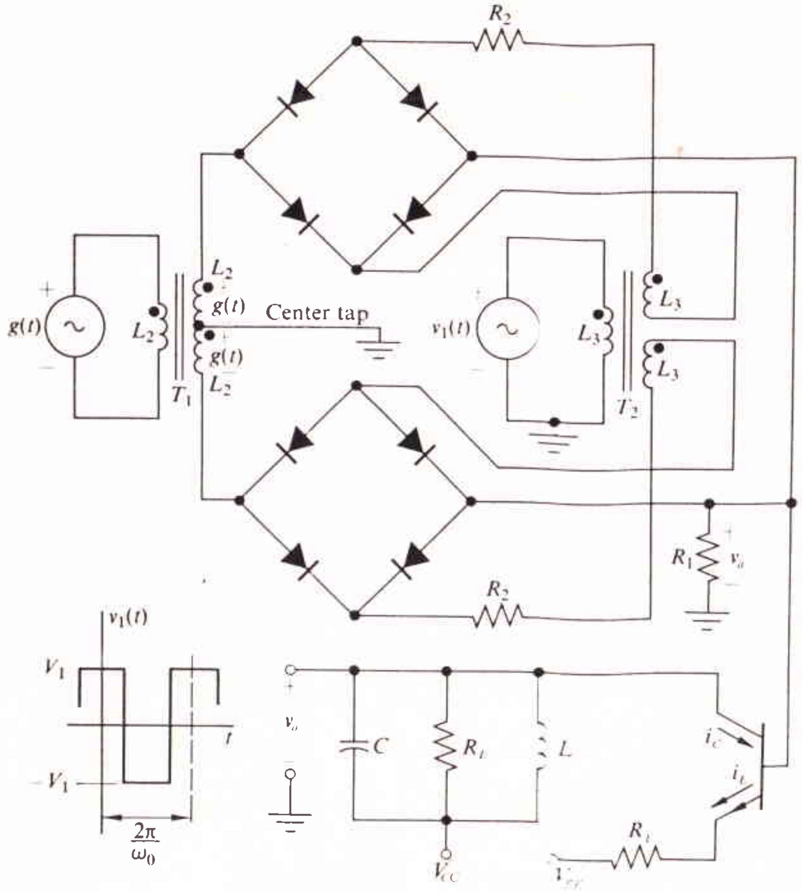


Fig. 8.4–7 Practical balanced chopper modulator.

closely coupled transformer with a midband capable of passing the main frequency components of $v_1(t)$. If $v_1(t)$ is a large-amplitude sine wave, then the restrictions on V_1 are nominal. [Although unity turns ratios are indicated for the two transformers, other turns ratios merely introduce a scale-factor change in $g(t)$ and $v_1(t)$.] If we assume that R_1 is not loaded by the transistor and that Eqs. (8.4–6) and (8.4–11) are satisfied, then $v_a(t)$ may be expressed as

$$v_a(t) = \frac{g(t)R_1}{R_1 + r_d}S'(t) \qquad (8.4\text{–}12)$$

where $S'(t)$ is given by Eq. (8.2–19). And if we assume that the transistor does not saturate, we may write

$$i_C = \alpha v_a(t)/R_E$$

d in turn

$$v_o(t) = V_{CC} - \left[\frac{4\alpha g(t)}{\pi R_E}\,\frac{R_1}{R_1 + r_d}\right] * z_{11_L}(t)\cos\omega_0 t \qquad (8.4\text{–}13)$$

here $z_{11_L}(t)$ is the low-pass equivalent impulse response of the output parallel RLC cuit. If the output filter passes $g(t)$ undistorted while removing the $3\omega_0$ component $v_o(t)$, then $v_o(t)$ reduces to the desired form

$$v_o(t) = V_{CC} - \frac{4\alpha R_L}{\pi R_E}\,\frac{R_1}{R_1 + r_d}g(t)\cos\omega_0 t. \qquad (8.4\text{–}14)$$

SQUARE-LAW MODULATOR

e square-law device, although quite attractive as a mixer, finds very little application an amplitude modulator. The basic reason for this is that most physical devices ve half-square-law characteristics rather than full-square-law characteristics. As saw in Section 8.2, unless a full-square-law characteristic exists, not only is supessed carrier modulation impossible but also normal AM with 100% modulation impossible. Consequently, unless a "quick and dirty" low-index modulator will isfy the requirements of the situation, the other modulators discussed in this apter are usually employed. Therefore, we shall look only briefly at one squarev modulator constructed with a junction FET operating within its saturation region.

A typical square-law FET modulator is shown in Fig. 8.5–1. If for this circuit assume that the FET operates within its saturation (square-law) region and that is much less than the output impedance of the FET, then we may approximate drain current as

$$i_D = I_{DSS}\left(1 - \frac{v_{GS}}{V_P}\right)^2, \qquad (8.5\text{–}1)$$

ere V_P is the pinch-off voltage and I_{DSS} is the drain current with $v_{GS} = 0$ and $= -V_P$. For the bias arrangement shown in Fig. 8.5–1, i_D reduces to

$$i_D(t) = I_{DSS}\left(\frac{v_1 + v_2}{V_P}\right)^2 = I_{DSS}\left(\frac{V_1\cos\omega_0 t + A[1 + mf(t)]}{V_P}\right)^2. \qquad (8.5\text{–}2)$$

ce the component of $i_D(t)$ centered about ω_0 is

$$\frac{2I_{DSS}V_1A}{V_P{}^2}[1 + mf(t)]\cos\omega_0 t,$$

n with the assumption that the output filter removes the low-frequency and secondmonic components of i_D, $v_o(t)$ is given by

$$v_o(t) = V_{DD} - \frac{2I_{DSS}V_1A}{V_P^2}[1 + mf(t)] * z_{11_L}(t)\cos\omega_0 t. \qquad (8.5\text{–}3)$$

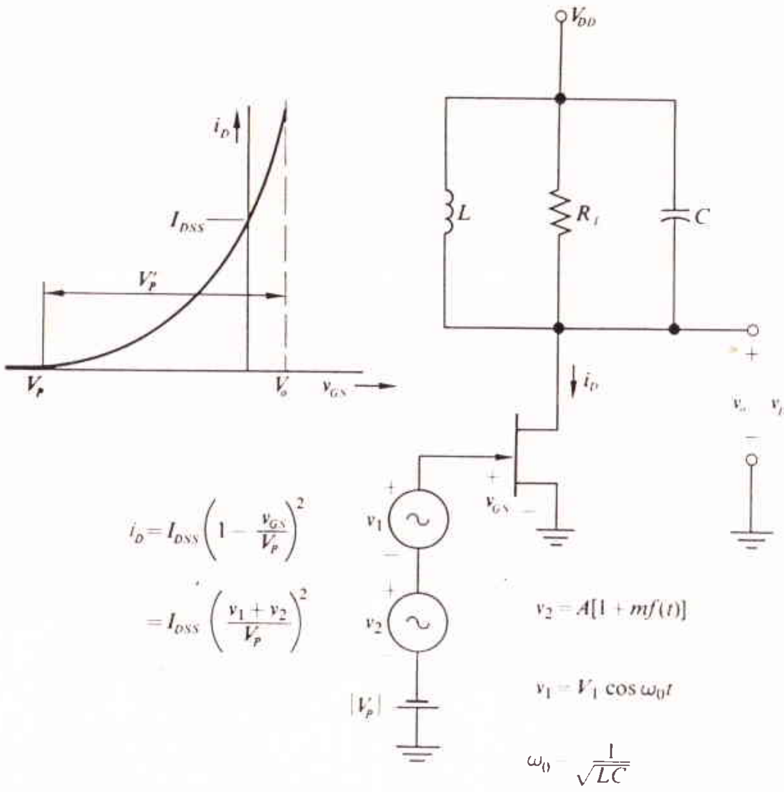

V_{DD}
i_D
I_{DSS}
V_P'
V_P
V_o
v_{GS}
L
R_L
C
i_D
+
$v_o = v_{DS}$
−
v_{GS}
$i_D = I_{DSS}\left(1 - \frac{v_{GS}}{V_P}\right)^2$
$= I_{DSS}\left(\frac{v_1 + v_2}{V_P}\right)^2$
v_1
v_2
$|V_P'|$
$v_2 = A[1 + mf(t)]$
$v_1 = V_1\cos\omega_0 t$
$\omega_0 = \frac{1}{\sqrt{LC}}$

Fig. 8.5–1 FET square-law modulator.

If, in addition, the output filter is flat over the band of frequencies occupied by the AM signal, $v_o(t)$ simplifies to the desired form

$$v_o(t) = V_{DD} - \frac{2I_{DSS}V_1AR_L}{V_P^2}[1 + mf(t)]\cos\omega_0 t. \qquad (8.5\text{–}4)$$

To realize $v_o(t)$ in the form of Eq. (8.5–4) we must restrict $v_o(t) > -|V_P|$ for all time to ensure that operation remains within the saturation region. This may be accomplished for any FET parameters by choosing a sufficiently small value of R_L.

In addition,

$$v_1(t) + v_2(t) \geq 0$$

for all time to keep the FET from being cut off and

$$v_1(t) + v_2(t) \leq |V_P| + V_0 = V'_P$$

for all time to keep the gate-to-source diode from turning on. These two restrictions ply that A, V_1, and m must satisfy the following inequalities:

$$
\begin{aligned}
A(1 - m) - V_1 &\geq 0, \\
A(1 + m) + V_1 &\leq V'_P.
\end{aligned}
\tag{8.5–5}
$$

If we desire the largest possible output with the smallest value of R_L (to ensure $_L \ll r_o$), then we satisfy Eq. (8.5–5) with an equality to obtain

$$A = \frac{V'_P}{2} \qquad \text{and} \qquad V_1 = \frac{V'_P}{2}(1 - m).$$

he output voltage given by Eq. (8.5–4) then reduces to

$$v_o(t) = (1 - m)\frac{I_{DSS}R_L}{2}\left(\frac{|V_P| + V_0}{V_P}\right)^2 [1 + mf(t)] \cos \omega_0 t. \tag{8.5–6}$$

Iere again we see explicitly that we can achieve a high modulation index only at the xpense of reducing $v_o(t)$ in amplitude or increasing R_L to the point where it is an ppreciable fraction of the FET output resistance r_o. If $R_L \rightarrow r_o$, then square-law peration no longer exists.

As a numerical example let

$$V_{DD} = 12\text{ V}, \qquad V_P = -4\text{ V}, \qquad V_0 = 0.7\text{ V}, \qquad I_{DSS} = -4\text{ mA},$$

$$\omega_0 = 10^7\text{ rad/sec}, \qquad \text{and} \qquad \omega_m = 2.5 \times 10^5\text{ rad/sec}.$$

[ω_m is the maximum frequency component of $f(t)$]. Let us now choose parameters for the circuit of Fig. 8.5–1 to achieve a maximum output voltage with $m = \frac{1}{2}$.

Since $|v_o(t)|_{\max} < V_{DD} + V_P$ to ensure saturation-region operation, then from Eq. (8.5–6) we obtain $R_L < 3.87\ \Omega$; thus a reasonable choice would be $R_L = 3.6\text{ k}\Omega$, which is much less than the typical values of r_o, say 100 kΩ. With this value R_L and a choice for the bandwidth of the output-tuned circuit of 5×10^5 rad/sec (the -3 dB point corresponds to ω_m),

$$C = \frac{1}{(3.6\text{ k}\Omega)(5 \times 10^5\text{ rad/sec})} = 566\text{ pF}$$

and in turn $L = 1/\omega_0^2 C = 18\ \mu\text{H}$. Finally, $A = 2.35$ V, $V_1 = 1.175$ V, and

$$v_o = (4.96\text{ V})[1 + mf(t)] \cos \omega_0 t. \tag{8.5–7}$$

The square-law modulator with the above parameter values is shown in Fig. 8.5–2. The gate bias circuitry is obtained by first noting from Fig. 8.5–1 that

$$
\begin{aligned}
v_{GS} &= V_P + A[1 + \tfrac{1}{2}f(t)] + V_1 \cos 10^7 t \\
&= (-1.65\text{ V}) + (1.175\text{ V})f(t) + (1.175\text{ V}) \cos 10^7 t
\end{aligned}
\tag{8.5–8}
$$

and then applying the -1.65 V to the gate by bleeding down the -12 V supply and coupling the ac sources to the gate through a capacitor.

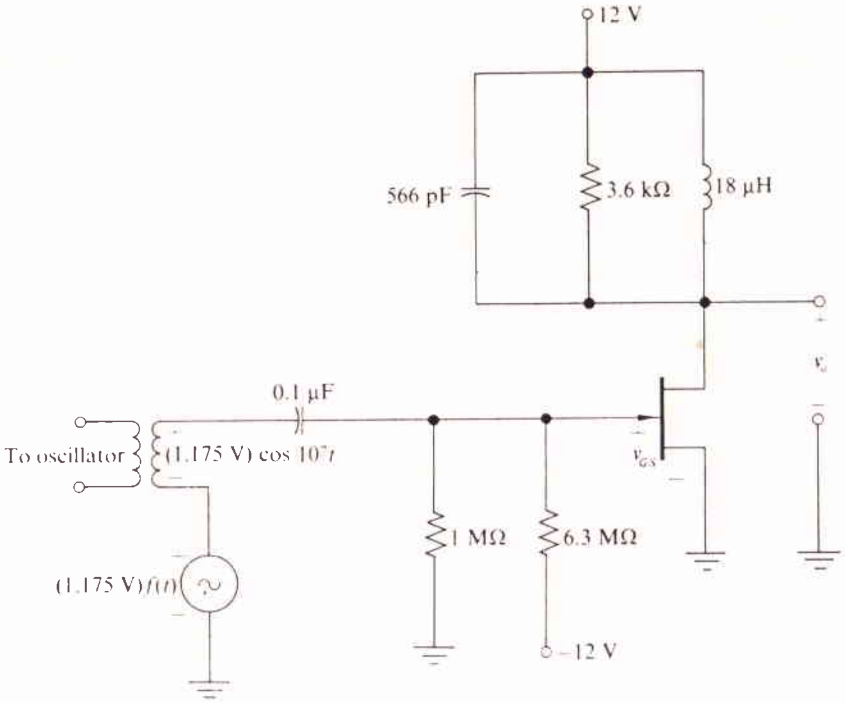


Fig. 8.5–2 FET modulator.

8.6 TUNED-CIRCUIT MODULATORS

As we shall see in Chapter 9, it is generally difficult to provide linear, high-power output amplifiers for AM signals; therefore, one usually tries to accomplish this modulation at as high a power level as is possible. In vacuum tube transmitters such modulation is almost inevitably performed in the last stage. To the extent that the transistor behaves as a current source, the triode vacuum-tube-type modulator cannot simply be "transistorized." As we shall see in Section 9.9, it is possible to construct efficient transistor power amplifiers if one both drives them with a modulated signal and simultaneously modulates the collector supply. It is the purpose of this section to present an idealized version of a circuit that can be used to perform the initial modulation of the carrier signal so that this premodulated signal may then be used to drive a power-amplifying stage. As we shall see, at first glance both the modulating and the power-amplifying stages look identical, but they do indeed operate in somewhat different fashions.

The basic modulating circuit being considered is shown in Fig. 8.6–1. The operation of this circuit requires the collector-base junction of the transistor to become saturated (or to turn on) at the peak of every driving carrier cycle. The current pulse that flows as a result of this saturation has two effects. The direct effect in the output circuit is to cause the output tuned-circuit voltage amplitude to follow the variations

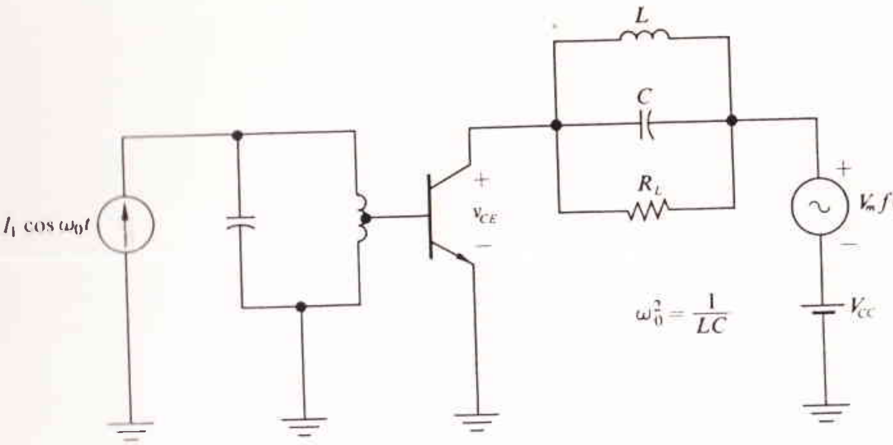
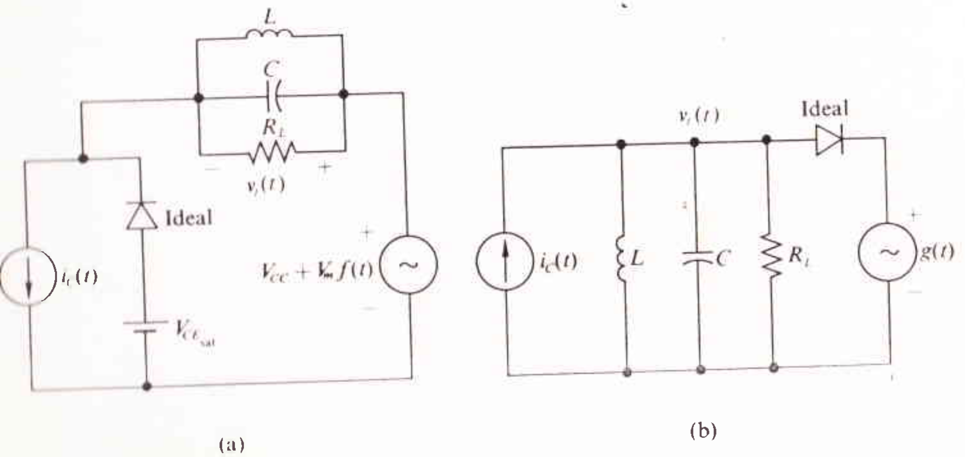


Fig. 8.6–1 Saturating collector AM modulator.

$f(t)$. The indirect or reflected effect is to increase the loading in the base circuit, hus effectively causing the input driving voltage also to follow the variations in $f(t)$.

In order to concentrate our attention on one problem at a time, we assume for he present that the input voltage is supplied by a voltage source. That is, we neglect ny possible variation of the input driving signal. Also, in order to be able to make ery simple calculations, we make the further assumption that the voltage drive is ufficiently large ($x \gg 10$) so that the transistor collector current may be assumed o flow as a train of impulses of strength q.

If the transistor is modeled as a current source in parallel with an ideal diode nd a battery $V_{CE_{sat}}$, then by splitting the current source into two parts [and then emoving the one in parallel with the $V_{CC} + V_m f(t)$ voltage generator] and redrawing



(a)

(b)

Fig. 8.6–2 Two equivalent forms for the model of the circuit of Fig. 8.6–1.

the circuit, one can reduce the circuit from the form of Fig. 8.6–2(a) to the form of Fig. 8.6–2(b). If m is defined as

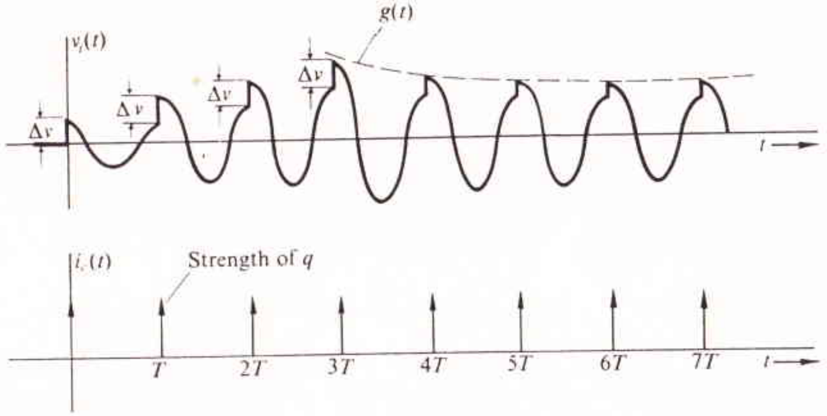
$$m = \frac{V_m}{V_{CC} - V_{CE_{sat}}}$$

and $g(t)$ is defined as

$$g(t) = (V_{CC} - V_{CE_{sat}})[1 + mf(t)],$$

then this circuit reduces to the form shown previously in Fig. 8.2–6.

The initial buildup of the tuned-circuit voltage in such a circuit is shown in Fig. 8.6–3. This buildup continues until the step on the tip of one of the carrier cycles exceeds $g(t)$ and the tuned-circuit voltage is caught at this value. Thereafter, so long as $g(t)$ does not vary too rapidly, the diode will conduct on every peak and the circuit will "ring" between the peaks at its natural frequency, ω_0.



$$i_c(t) = q\sum_{n=0}^{\infty} \delta(t - nT), \quad T = 2\pi/\omega_0$$

Fig. 8.6–3 Transient buildup of the tank voltage in a tuned-circuit modulator.

If the tuned-circuit Q_T is high enough (say more than 30) so that the energy decay between resettings can be considered linear, then for any constant value of $g(t)$ the size of the step at each peak will be approximately $g(t)\pi/Q_T$. Therefore, the higher the value of Q_T, the smaller the distortion at each cycle peak. In a practical circuit we do not drive with impulses; hence the output voltage does not have steps. For example, with current pulses that are 90° wide the voltage transient in the tuned-circuit voltage will be nearly imperceptible, even though the fundamental current into the tuned circuit is essentially the same in the case of impulse, and in the case of wider current pulses.

Failure-to-Follow Distortion

If $g(t)$ increases too rapidly in a cycle, then the maximum step in the tuned-circuit voltage, $\Delta V = q/C$, will be insufficient to cause the diode to conduct and the tuned-circuit voltage will not track the modulating signal.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE [illegible]

It can be shown for the impulse drive case that the necessary condition to prevent ward-going failure-to-follow distortion is the satisfaction of the condition

$$[\dot{g}(t_0)/\alpha] + g(t_0) \leq I_1 R_L, \tag{8.6–1}$$

ere α is the real part of the tuned-circuit pole, t_0 is the time at any cycle peak, and $= 2q/T$ is the fundamental component of the collector impulse train. A simple oof of Eq. (8.6–1) requires only (1) the assumption of a high enough value for Q_T at the envelope decay during a cycle is linear and (2) the assumption of a slow rate change for $g(t)$. Further investigation indicates that the result is valid for all values Q_T and of $\dot{g}(t_0)$.

Intuitively it is obvious that the input drive must always be such that $I_1 R_L$ exceeds e peak value of $g(t)$, otherwise, failure-to-follow distortion will certainly result. quation (8.6–1) establishes a further relationship among the drive signal, the circuit. d the rate of change of the modulation.

hase Distortion

$g(t)$ decreases more rapidly than the natural envelope decay of the tuned-circuit oltage between current pulses, then the period between diode conductions will fall elow $T = 2\pi/\omega_0$ and phase modulation of the carrier component of the tuned-ircuit voltage will result. To prevent such distortion,

$$g(t_0 + T) \geq g(t_0)e^{-\alpha T} \tag{8.6–2}$$

s a necessary and sufficient condition.

When Q_T is high and the rate of change of $g(t)$ is low, this equation may be approxi-mated by

$$[\dot{g}(t_0)/\alpha] + g(t_0) \geq 0. \tag{8.6–3}$$

As in the case of failure-to-follow distortion, a more involved analysis shows that Eq. (8.6–3) is in fact valid without restrictions on either $\dot{g}(t)$ or Q_T.

To prevent either type of distortion from occurring, the two previous restrictions may be combined to yield

$$I_1/2C \geq \dot{g}(t_0) + \alpha g(t_0) \geq 0, \tag{8.6–4}$$

where excessive positive values for $\dot{g}(t_0)$ lead to the violation of the left-hand inequality and excessive negative values of $\dot{g}(t_0)$ lead to the violation of the right-hand inequality.

Sinusoidal Modulation

For the specific case of sinusoidal modulation where

$$g(t) = (V_{CC} - V_{CE_{\text{sat}}})(1 + m \cos \omega_m t),$$

one may evaluate Eq. (8.6–4) in order to show the restrictions necessary on m to avoid either type of distortion. For downward-going modulation the restriction is

$$m \leq \frac{1}{\sqrt{1 + (\omega_m/\alpha)^2}}; \tag{8.6–5}$$

for upward-going modulation the restriction is

$$m \leq \frac{[I_1 R_L/(V_{CC} - V_{CE_{\text{sat}}})] - 1}{\sqrt{1 + (\omega_m/\alpha)^2}}. \tag{8.6–6}$$

Hence if $I_1 R_L$ is greater than $2(V_{CC} - V_{CE_{\text{sat}}})$, then the downward-going restriction governs in all cases. For 100% modulation this restriction on the minimum size for I_1 is imposed at any rate; hence no additional hardship is added by the rate of variation of the modulation.

One way to ensure that m always satisfies the necessary restrictions is to place the modulating signal through a low-pass filter with a −3 dB bandwidth lower than α and to obtain $g(t)$ at the filter output. Specifically, if the filter transfer function is $H(j\omega)$, then

$$g(t) = (V_{CC} - V_{CE_{\text{sat}}})H(0)\left\{1 + \frac{|H(j\omega_m)|}{H(0)} \cos\left[\omega_m t + \arg H(j\omega_m)\right]\right\}. \tag{8.6–7}$$

Since the resultant value of m from Eq. (8.6–7) is $|H(j\omega_m)|/H(0)$, then, if this function is always equal to or less than $1/\sqrt{1 + (\omega_m/\alpha)^2}$, Eq. (8.6–5) will always be satisfied. Figure 8.6–4 illustrates this situation.

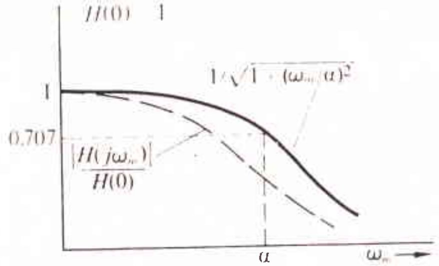


Fig. 8.6–4 Prefiltering of the modulation to ensure no distortion in the modulated output $[I_1 R_L \geq 2(V_{CC} - V_{CE_{\text{sat}}})]$.

For a practical circuit the phase-distortion restriction should be reduced, since with a spread-out current pulse an earlier conduction time for the diode conduction should lead to a less abrupt phase shift in the output. Certainly, if Eq. (8.6–5) is satisfied, then distortion should be avoided.

To summarize: We have seen that, if the collector-base diode of a transistor conducts on every cycle, then the circuit of Fig. 8.6–1 performs as an AM modulator. As was pointed out in a previous section, it is possible to obtain 100% modulation from such a circuit without the strain of driving the tuned circuit between zero and $2(V_{CC} - V_{CE_{\text{sat}}})$ if one merely modulates to a peak value of, say, 70% and then subtracts a carrier component sufficient to bring the valleys down to zero.

Once again we shall point out that, though the modulator of this section looks e the modulated power amplifier of Section 9.9 the operation of the two devices is t identical. The modulated power amplifier can operate perfectly satisfactorily en if the collector-base diode never goes into conduction, whereas the whole basis operation for the modulator of this section is the conduction of this diode.

ROBLEMS

1 Find the amplitude and frequency of the spectral components of $v(t) = g(t) \cos \omega_0 t$ for each of the following cases:

a) $g(t) = (5\text{ V}) \cos \omega_m t$, $\omega_m \ll \omega_0$.

b) $g(t) = (5\text{ V}) + (3\text{ V}) \cos \omega_m t$, $\omega_m \ll \omega_0$.

c) $g(t)$ is a symmetric square wave which varies between 0 V and 5 V with a period $T = 1/\omega_m \gg 1/\omega_0$.

2 The voltage $v(t)$ given in Problem 8.1 appears at the input of an antenna having an input resistance of 100 Ω in the neighborhood of ω_0. Determine the carrier power and the sideband power delivered to the antenna for each $g(t)$ given in Problem 8.1.

3 The voltage $v(t)$ given in Problem 8.1 is applied to a single tuned circuit having a center frequency of ω_0, a bandwidth of $2\omega_m$, and unity transmission at ω_0. Find the voltage at the filter output for each $g(t)$. Sketch the frequency spectrum in each case.

4 The modulation $g(t) = (5\text{ V}) \cos \omega_1 t + (10\text{ V}) \cos \omega_2 t$ $(\omega_1 < \omega_2 \ll \omega_0)$ is applied to the SSB modulator shown in Fig. 8.1–5. Determine

a) $\hat{g}(t)$, the Hilbert transform of $g(t)$,

b) an expression for $v_{SSB}(t)$,

c) the spectral components of $v_{SSB}(t)$, and

d) the power $v_{SSB}(t)$ delivers to a 100 Ω resistor when placed across it.

5 Show that, when $g(t)$ is placed through two Hilbert transformers in cascade, the output of the second Hilbert transformer is $-g(t)$; that is, $H[Hg(t)] = -g(t)$.

6 Determine the amplitude of the carrier which must be subtracted from $v(t) = (10\text{ V})[1 + 0.2f(t)] \cos \omega_0 t$ to yield a modulation index of 0.9.

8.7 In the circuit shown in Fig. 8.2–2, $g(t) = (5\text{ V})(1 + \cos 10^4 t)$, $\omega_0 = 10^7$ rad/sec, and $H(j\omega)$ is a single tuned filter having a bandwidth of 2×10^4 rad/sec and a center frequency of 3×10^7 rad/sec. Assuming that the filter has unity transmission at resonance, find an expression for $v_o(t)$.

8.8 In the circuit shown in Fig. 8.2–5, $A = 10$ V, $V_1 = 5$ V, and

$$v_a = \begin{cases} 0, & v_i \le 0, \\ v_i + 4v_i^2, & v_i > 0, \end{cases}$$

where v_a and v_i are in volts. Calculate the maximum output modulation index which can be achieved without causing envelope distortion in $v_o(t)$.

8.9 For the circuit shown in Fig. 8.P–1, find $v_o(t)$. What is the modulation index? Sketch $v_o(t)$ if $f(t) = \cos 10^3 t$.

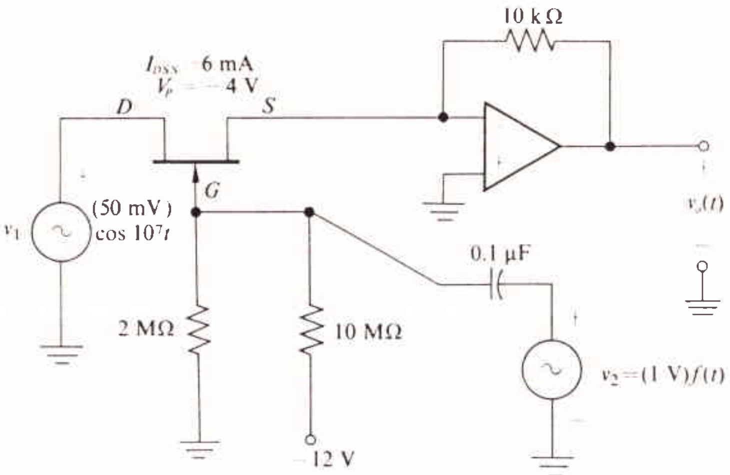


Figure 8.P–1

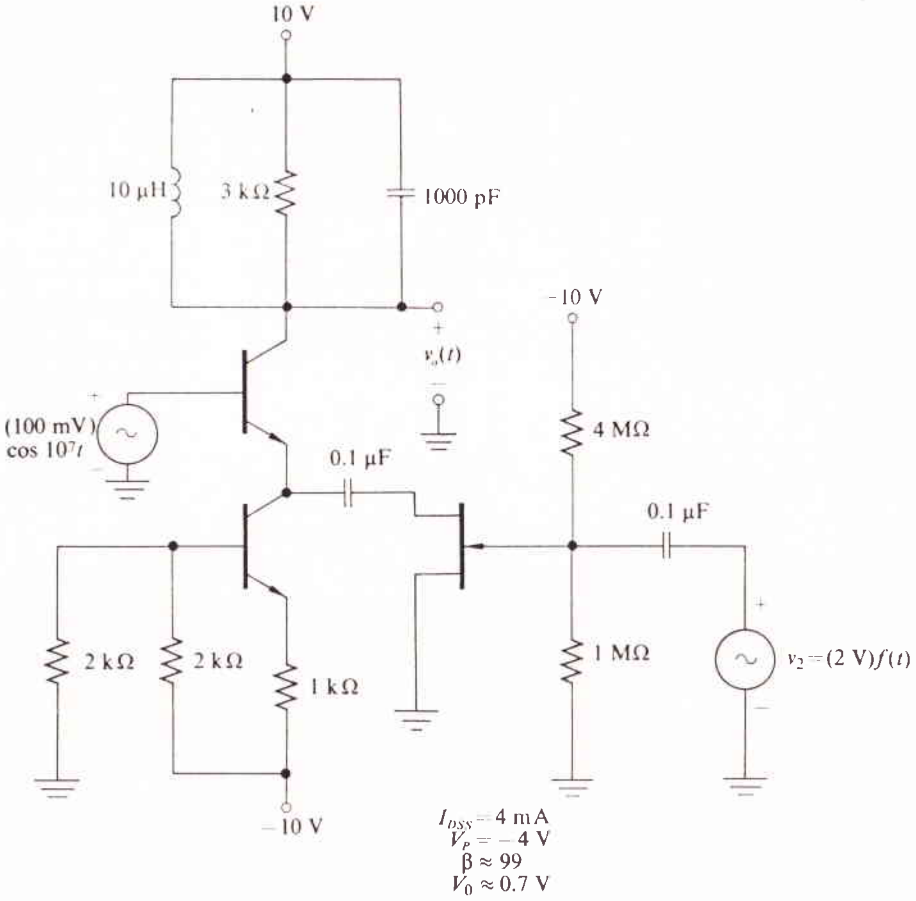


Figure 8.P–2

For the circuit shown in Fig. 8.P–2, calculate $v_o(t)$ if $f(t) = \cos 1.66 \times 10^5 t$. Repeat, assuming that $f(t)$ is a square wave with transitions between +1 and −1 and with a 100 μsec period.

a) For the circuit shown in Fig. 8.P–3, determine an expression for $v_o(t)$ assuming that none of the transistors saturate, that all the transistors are identical, and that

$$|v_1(t)| \le 100\text{ mV} \qquad \text{and} \qquad 0 \le v_2(t) \le |V_P| + V_0.$$

b) Evaluate $v_o(t)$ for the case where

$$I_{DSS} = 4\text{ mA}, \quad V_P = -2\text{ V}, \quad R_L = 2\text{ k}\Omega, \quad \pm V_{CC} = \pm 10\text{ V},$$

$$R_B = 3.9\text{ k}\Omega, \quad v_1 = 100\text{ mV} \cos 10^7 t, \quad \text{and} \quad v_2(t) = (1\text{ V})(1 + 0.6 \cos 10^3 t).$$

Does any transistor saturate?

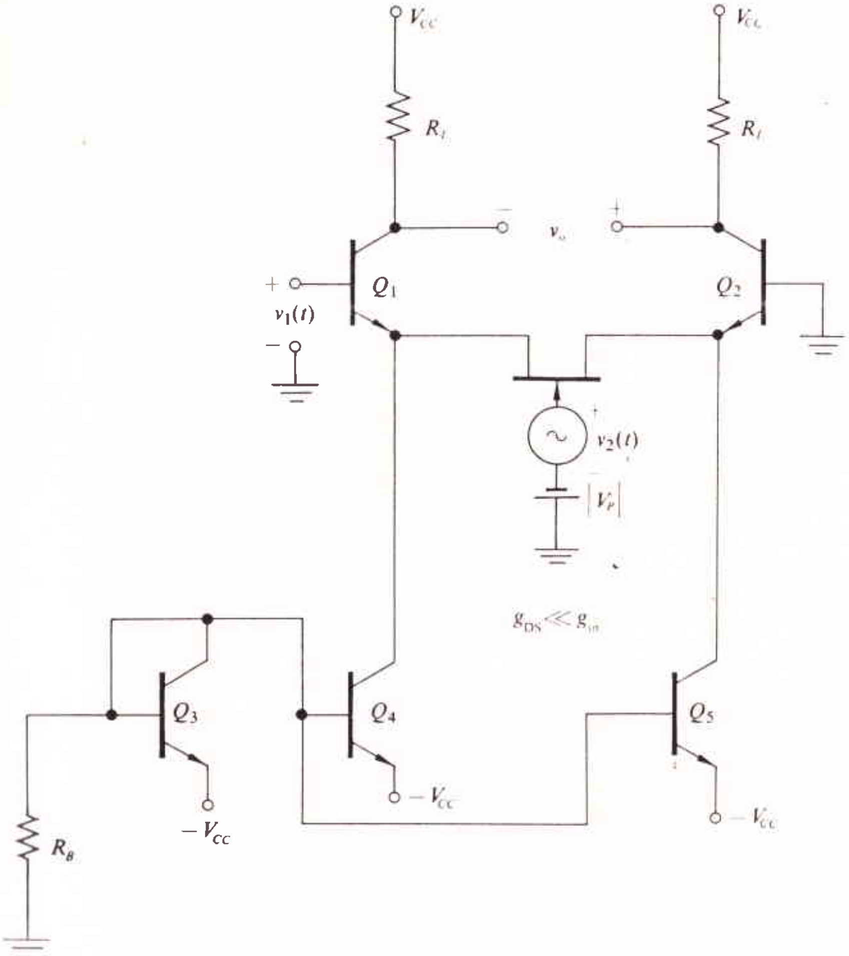


Figure 8.P–3

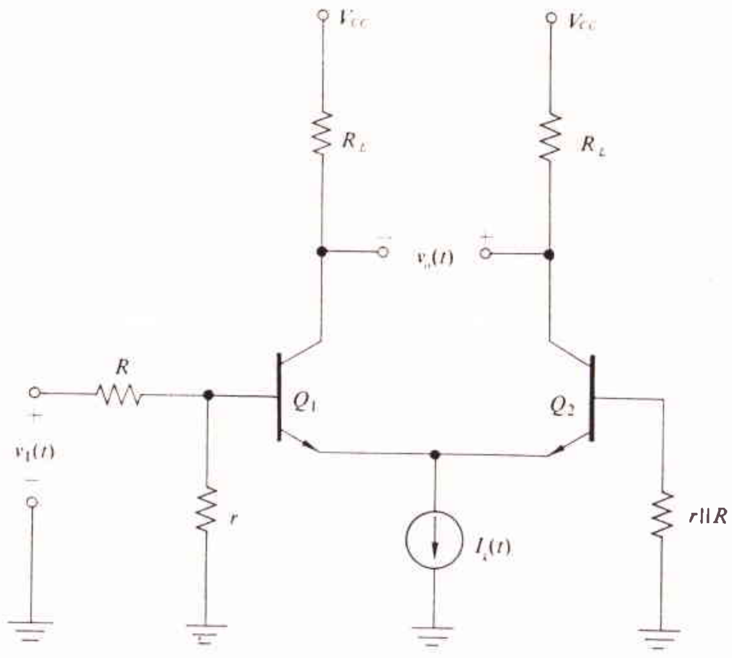


Figure 8.P–4

8.12 For the two-quadrant multiplier shown in Fig. 8.P–4, show that

$$v_o(t) = \frac{\alpha q R_L r}{2kT(R + r)} v_1(t) I_k(t),$$

provided that

a) Q_1 and Q_2 are identical with sufficiently high values of β so as not to load r,

b) neither Q_1 nor Q_2 saturates,

c) $I_k(t) \ge 0$, and

d) $|v_1 r/(R + r)| < 16$ mV.

Given

$$R = 1\text{ k}\Omega, \quad r = 10\ \Omega, \quad R_L = 2\text{ k}\Omega, \quad V_{CC} = 10\text{ V},$$

$$v_1(t) = 1\text{ V} \cos 10^7 t, \quad \text{and} \quad I_k(t) = (2.5\text{ mA})[1 + mf(t)],$$

find an expression for $v_o(t)$. Are the assumptions justified?

8.13 Figure 8.P–5 illustrates a circuit for unbalancing a differential output. Show that if Q_i and Q_i' are identical transistors, $i = 1, 2, 3$, with high values of β, then

$$v_o \approx (i_2 - i_1) R_L.$$

(This is the same voltage that would be obtained differentially between the collectors of Q_5 and Q_4 if R_L were placed in the collectors of both Q_5 and Q_4.)

Show how this unbalancing network may be used in conjunction with the circuits of Figs. 8.3–10, 8.P–3, and 8.P–4. Write an expression for $v_o(t)$ for each case.

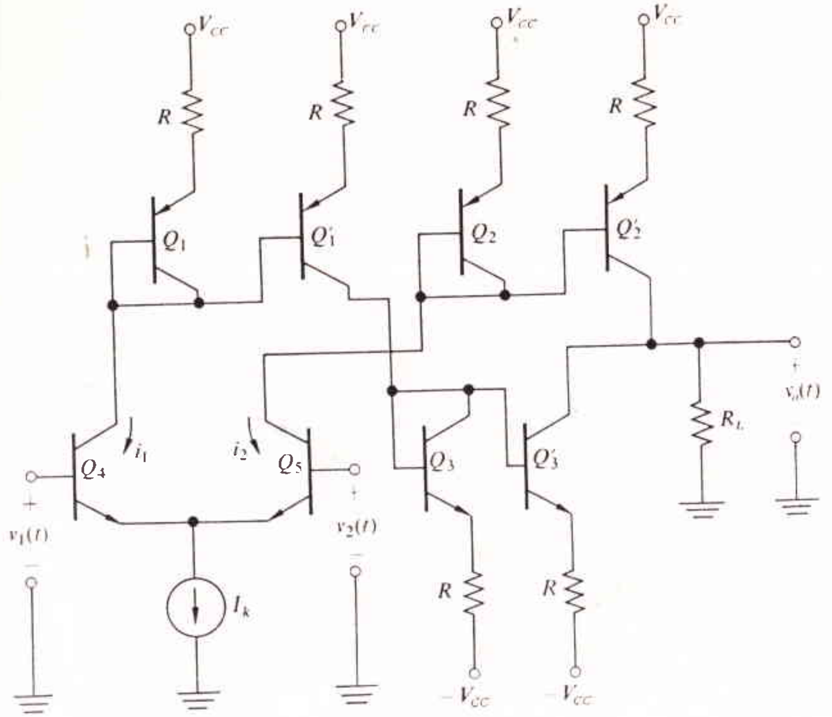


Figure 8.P–5

14 For the circuit shown in Fig. 8.P–6, determine an expression for $v_o(t)$ as well as $v_{o1}(t)$ and $v_{o2}(t)$. Assume that all transistors are identical with $\beta > 200$ and $V_{BE} \approx 0.7$ V.

15 Determine an expression for $v_o(t)$ in the circuit shown in Fig. 8.P–7. Assume that $\beta = 99$, $V_{BE} \approx 0.7$ V, and that the tuned circuit passes the modulation.

16 For the circuit shown in Fig. 8.4–1,

$$g(t) = (10\text{ V})\cos 10^3 t, \qquad \omega_0 = 10^7\text{ rad/sec}, \qquad R_1 = 1\text{ k}\Omega, \qquad R_L = 1\text{ k}\Omega,$$

$$C = 1000\text{ pF}, \qquad L = 10\ \mu\text{H}, \qquad \text{and} \qquad g_m = 10{,}000\ \mu\mho.$$

Determine the minimum value of V_1 and the maximum value of R_2 which permit proper circuit operation. With these values write an expression for $v_o(t)$.

17 For the circuit shown in Fig. 8.P–8, calculate $v_o(t)$ if $I_{DSS} = 4$ mA and $V_P = -4$ V.

18 Derive Eqs. (8.6–1) and (8.6–3).

19 For the circuit shown in Fig. 8.P–9,

$$i_i(t) = \sum_{n=-\infty}^{\infty} Q\delta(t - nT),$$

where $T = 2\pi \times 10^{-7}$ sec and $Q = 30\pi$ pC. Make an accurate sketch of $v_o(t)$. Indicate regions where failure-to-follow distortion occurs.

8.20 For the circuit of Problem 8.19, $v(t) = (10\text{ V})(1 + m\cos 10^5 t)$. Find the maximum value of m before either phase distortion or failure-to-follow distortion occurs.

8.21 For the circuit of Problem 8.19, $v(t) = [10 - 9\mu(t - t_i)]$ V. Assuming that t_i occurs on a negative peak of $v_o(t)$, write an expression for (or plot) $v_o(t)$ for two carrier cycles beyond t_i. Is there any phase distortion?

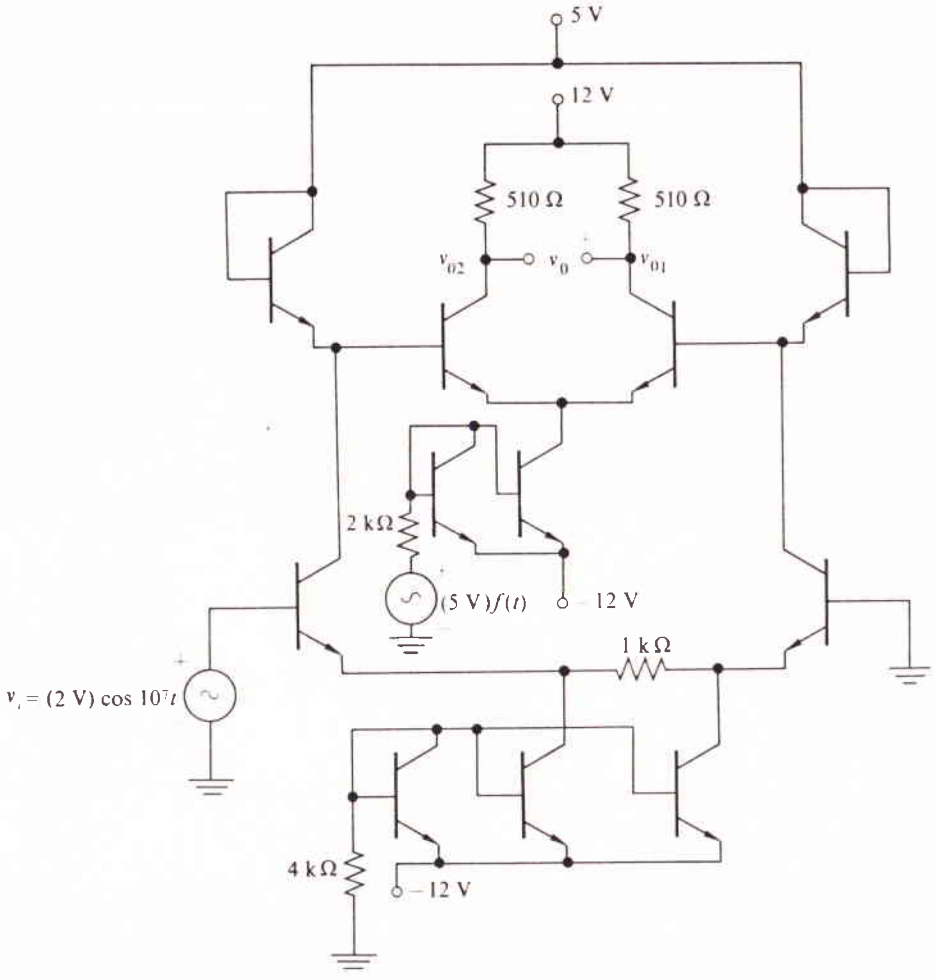


Figure 8.P–6

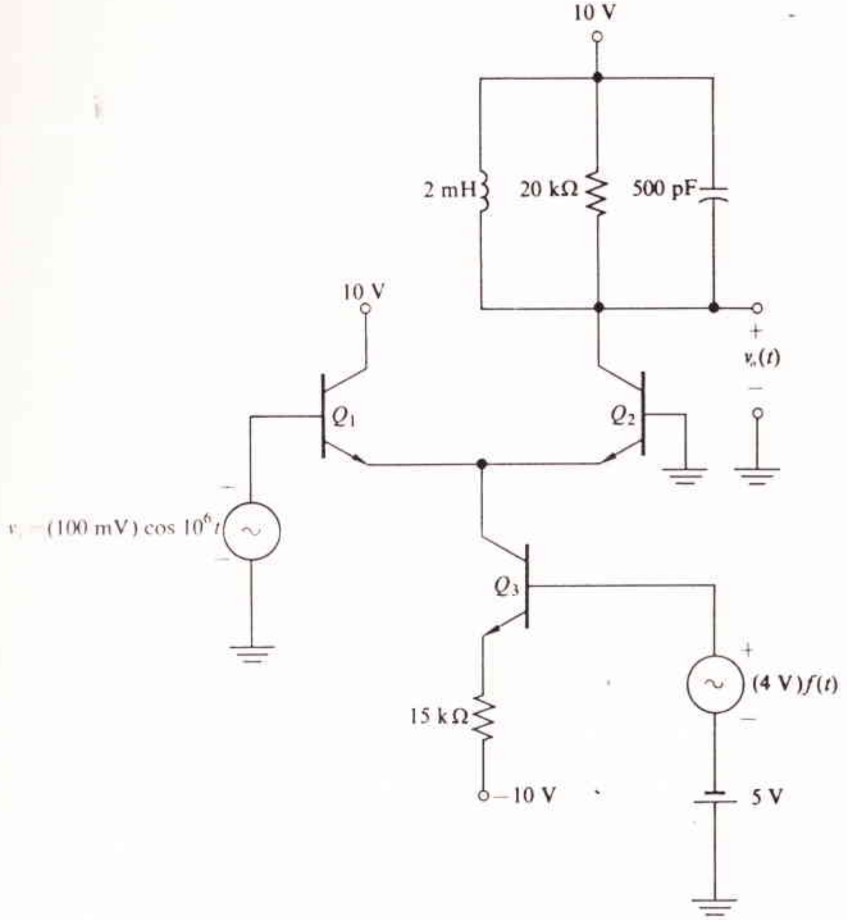


Figure 8.P–7

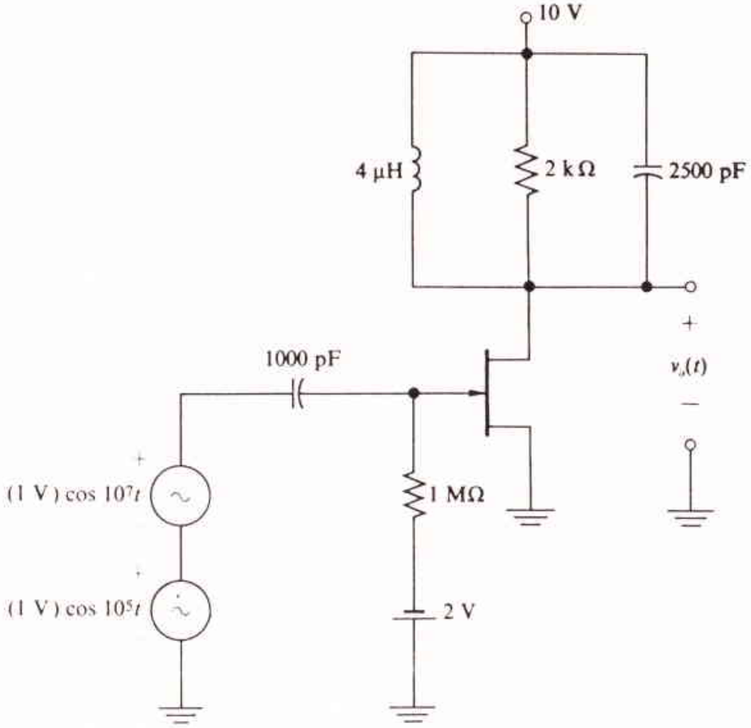


Figure 8.P–8

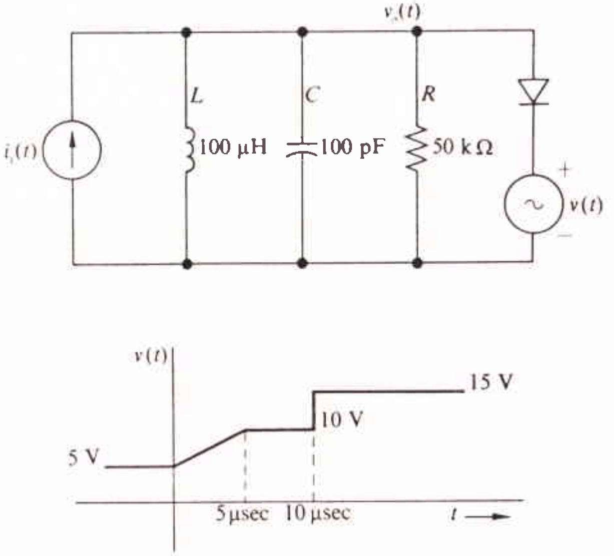


Figure 8.P–9

HAPTER 9

POWER AMPLIFIERS

previous chapters we have been concerned with performing some operation upon a nal, rather than with the efficient conversion of supply power into signal power. this chapter we shall concentrate upon this power conversion problem. Again we ll begin by considering the basic problems in a device-free context, and then we ll modify these general results by taking into account the additional limitations posed by specific types of devices.

Intuitively we see that the efficient conversion of supply power into signal power quires that the losses in the conversion device be minimized. If we consider the quency ranges and circuit adjustments where device voltage drops and device rrents are in phase, then it follows that device power is minimized if maximum rrent flows during minimum voltage drops while maximum voltage drops are companied by minimum current flows.

In addition, if the signal to be amplified is either absent or small for a substantial rt of the time (speech or music, for example), then it is clear that the long-term erage device dissipation will be minimized by minimizing the "standby" or Q-point wer through the device.

Power amplifier output for a given device is normally limited by one of two tors:

) excessive distortion or
) some device and/or circuit limitation (other than distortion).

Examples of the device limitations include: (a) voltage breakdown, (b) current handling limit, and (c) maximum allowable device temperature. Examples of combined circuit and device limitations include: (a) operating path limitations (e.g., violation of the "safe operating region" or SOAR†) and (b) thermal runway.‡

---

device manufacturer may specify regions of the characteristics of a certain device in which it is fe" to operate. These regions may be functions of the type of output circuitry and the shape frequency of the driving waveshape. Most high-power transistors have such regions as a tion of their specifications. Part of the design procedure for any power amplifier is to ensure t the operating path (see Fig. 9.4–3, for example) lies completely within the "safe" operating ion.

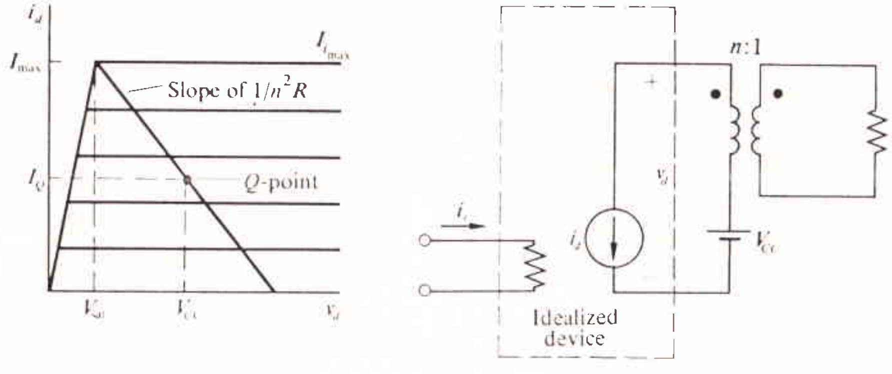
hermal runaway is a phenomenon wherein device self-heating causes more dc current to ; this leads to more device heating, and the process continues until the device is destroyed. s positive feedback problem is peculiar to transformer- (or choke-) coupled circuits, since h dc resistive loads the increased device current must decrease the device dc voltage and ntually the circuit will reach equilibrium. The problem is controlled by a combination of

9.1 "IDEAL" POWER AMPLIFIERS—CLASS A, SINGLE-ENDED

Initially we shall concentrate on the limitations imposed by distortion caused by the drive waveshape and the circuit configuration. That is, we neglect gradual internal device distortions by assuming that until the device reaches cutoff or saturation the output parameter (current or voltage) is a faithful copy of an input parameter (current or voltage).

We further simplify the situation by assuming that our devices are "current-source-like" or "voltage-source-like." For example, with a "current-source-like" device the output current depends only on the input parameters and is independent of the output voltage.

Figure 9.1–1 shows a broadband version of such an idealized device together with a possible circuit for its use as a power amplifier. If such a circuit is operated in such a way that i_d always exceeds, zero, i.e., the device never reaches cutoff, it is called a Class A amplifier.†



(a) (b)

Fig. 9.1–1 Device characteristic and possible Class A power amplifier circuit. (a) i_i as a parameter. (b) Idealized circuit.

---

reduced thermal resistances, increased bias stability, and the use of external temperature compensation elements. For details, see Chapter 3 of *Transistor Circuit Analysis* by M. V. Joyce and K. K. Clarke (Addison-Wesley, Reading, Mass., 1961).

† Power amplifier classes are defined in terms of the device conduction angle and/or the type of device operation:

Class A:	"Linear" operation, 360° conduction angle
Class B:	"Linear" operation, 180° conduction angle
Class C:	Fixed drive, less than 180° conduction angle
Class D:	Switched operation, conduction angle may vary with time from 0° to 360° or may be fixed
Class AB:	"Linear" operation, conduction angle less than 360° but more than 180°

In certain cases a particular circuit may be viewed as belonging simultaneously to several different classes.

It is apparent that among the variables to be chosen in this circuit are V_{CC}, n, R_L, e Q-point, the waveshapes of i_i, and the peak value of the driving waveshape. The oice of these variables indicated in Fig. 9.1–1 is the optimum one assuming a given $_C$, a given $I_{i_{\max}}$, and the desire to operate in a Class A fashion with the maximum ossible efficiency: i.e. $v_d = V_{\min} = V_{\text{sat}}$ when $i_d = I_{\max}$.

If i_i is symmetrical about $I_{i_{\max}}/2$ and never passes cutoff or $I_{i_{\max}}$, then the output rrent will be symmetrical about I_Q, which will be equal to I_{dc}. At the same time, nce we are assuming midband operation where the transformer reactances may be glected and the upper and lower breakpoints are widely separated, the device ltage will be swinging symmetrically about V_{CC}.

In some cases, certainly with very low-frequency signals or with devices that have ry short thermal time constants, one wishes to consider instantaneous rather than erage device power. However, it normally suffices to average the power over a cle of the lowest-frequency signal to be amplified. To locate the circuit of Fig. 1–1 in the power amplifier hierarchy, let us begin by calculating its power output, ower input, power dissipation, and power conversion efficiency η, all as a function of put waveshape.

$$
\begin{aligned}
&i_d = I_Q + I_1 \cos \omega t, \qquad I_Q + I_1 \le I_{\max},\\
&I_Q = I_{\text{dc}}, \qquad R_L = n^2 R,\\
&v_d = V_{CC} + I_1 R_L \cos \omega t, \qquad I_1 R_L \le V_{CC} - V_{\text{sat}},\\
&P_{\text{dc}} = I_{\text{dc}} V_{CC}, \qquad P_{\text{ac}} = \frac{I_1^2 R_L}{2}, \qquad \eta = \frac{I_1^2 R_L}{2 I_{\text{dc}} V_{CC}} = \frac{P_{\text{ac}}}{P_{\text{dc}}},
\end{aligned}
\tag{9.1–1}
$$

nd

$$
P_{\text{device}} = P_{\text{dc}} - P_{\text{ac}} = P_{\text{dc}}(1 - \eta) = P_{\text{ac}} \frac{1 - \eta}{\eta}, \tag{9.1–2}
$$

here P_{dc} is the power supplied by the battery and P_{ac} is the power consumed by the ad. (Actually P_{ac} is the power available at the input terminals of the transformer. ome small-output power transformers have transfer efficiencies as low as 0.60; ence all of P_{ac} may not actually reach the load.)

When this circuit has no sinusoidal drive, all the power supplied by the battery by the power supply must be dissipated by the device. As the drive increases, so oes η, and the device dissipation falls. With the maximum drive allowable before itoff or saturation occurs and gross distortion results, one finds

$$
\begin{aligned}
&P_{\text{dc}_{\max}} = I_{\text{dc}} V_{CC}, \qquad P_{\text{ac}_{\max}} = \frac{I_{\text{dc}}(V_{CC} - V_{\text{sat}})}{2},\\
&\eta_{\max} = \frac{V_{CC} - V_{\text{sat}}}{2 V_{CC}},\\
&P_{\text{device}_{\min}} = \frac{(V_{CC} + V_{\text{sat}}) I_{\text{dc}}}{2},
\end{aligned}
\tag{9.1–3}
$$

here $I_1 R_L = V_{CC} - V_{\text{sat}}$ and $I_1 = I_{\text{dc}}$. Suppose that $V_{\text{sat}} = 0.05 V_{CC}$; then $\eta_{\max} = 0.475$

while P_{device} falls from $V_{CC} I_{\text{dc}}$ with no sinusoidal drive to $0.525 V_{CC} I_{\text{dc}}$ with full sinusoidal drive and P_{ac} climbs from zero with no drive to $0.475 V_{CC} I_{\text{dc}}$ at full drive.

Figure 9.1–2 indicates the relative values of P_{dc}, P_{ac}, P_{device}, and η as a function of I_1/I_{dc} or $V_1/(V_{CC} - V_{\text{sat}})$. In all cases the normalization is with respect to $P_{\text{dc}_{\max}}$.

From Fig. 9.1–2 it is obvious that, if the drive signal varies in amplitude and is occasionally zero, then the device must be capable of dissipating $I_{\text{dc}} V_{CC}$ watts and the overall long-term efficiency may be 10% or less.

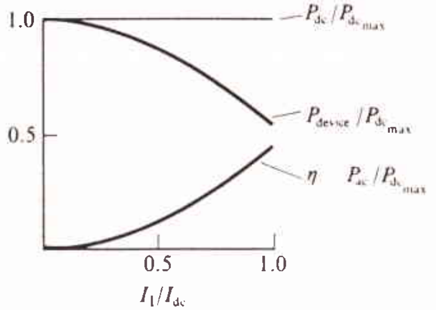


Fig. 9.1–2 Normalized powers and efficiency for a sinusoidally driven Class A amplifier ($V_{\text{sat}} = 0.05 V_{CC}$).

Square-Wave Drive—Broadband Load

If we assume that the device and the output transformer will both handle square waves without distortion, then, if the peak-to-peak value of the square-wave current is I_p, where $I_p \le 2 I_{\text{dc}}$, it follows that

$$
P_{\text{ac}} = \frac{I_p^2 R_L}{4}, \qquad \eta = \frac{I_p^2 R_L}{4 I_{\text{dc}} V_{CC}} = \left(\frac{I_p}{2 I_{\text{dc}}}\right)^2 \frac{V_{CC} - V_{\text{sat}}}{V_{CC}}. \tag{9.1–4}
$$

In this case $\eta_{\max} = 1 - V_{\text{sat}}/V_{CC}$, while P_{device} decreases from P_{dc} with no drive to $I_{\text{dc}} V_{\text{sat}}$ at maximum drive.

Figure 9.1–3 indicates the normalized variation as a function of $I_p/2I_{\text{dc}}$ or $V_p/(V_{CC} - V_{\text{sat}})$. If the drive ever drops to half its maximum value, the efficiency will

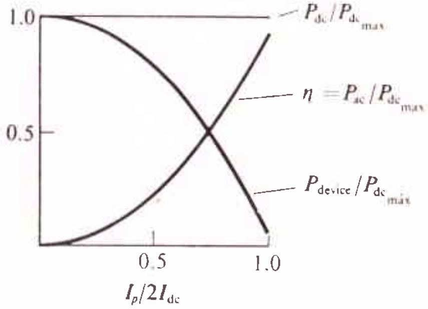


Fig. 9.1–3 Normalized powers and efficiency for a square-wave-driven Class A amplifier ($V_{\text{sat}} = 0.05 V_{CC}$).

drop from 95% to 24% while the device dissipation will rise from 50% of the dc Q-point power to 76% of this power. Hence, except for the special case of continuous maximum drive, the device must be prepared to dissipate $I_{dc}V_{CC}$ and the average overall efficiency, while higher than in the sine-wave case, may easily be 20% or less.

Class A Narrowband Operation

As we have seen, a power amplifier consists of a controlling device, a load, a basic power source, and a driving waveform. We have also seen that the results obtained with power amplifiers are quite sensitive to the type of driving waveshape. The results are also a function of the load circuit. In general, there are two useful types of loads. In the broadband case one wishes the load to look "resistive" and "constant" over a frequency range of at least 10/1 and more normally 100/1 or even 1000/1 or more. In the narrowband or high-Q case one is often satisfied if the 70% variation points of the load impedance are within several percent of the center frequency. Again one would normally like the impedance to be resistive at the center frequency.

One basic difference between the two types of loads is that in the broadband case the load cannot be counted on to remove signal distortion, while in the narrowband case the filtering properties of the load may lead to output voltages or currents that are quite different from the driving waveshapes.

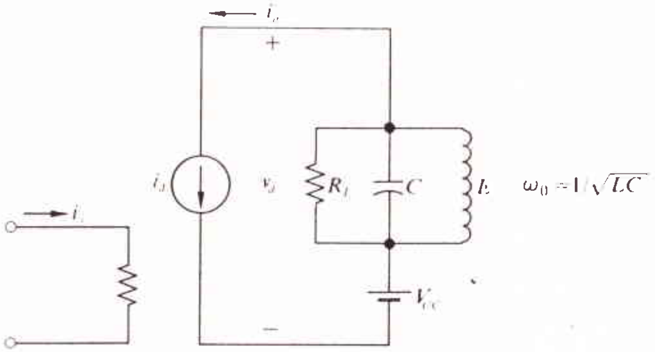
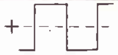


Fig. 9.1–4 Class A narrowband amplifier.

If we replace the broadband transformer and load combination by a high-Q parallel RLC circuit, then the circuit of Fig. 9.1–4 will result. In order to see the different effects of driving waveshapes and to compare the broadband and narrowband cases, let us calculate the results for this circuit with sine- and square-wave output device currents. (Again we assume that i_d contains a dc component of half its maximum value, I_{max}, $R_{\text{sine wave}} = 2(V_{CC} - V_{sat})/I_{max}$ while $R_{\text{square wave}}$ equals $(V_{CC} - V_{sat})/2I_{max}$. We also assume that the driving waveshape has a fundamental radian frequency of ω_0.) The results are shown in Table 9.1–1.

The results for the case of sinusoidal drive are identical to those for the broadband case as summarized in Fig. 9.1–2. On the other hand, the case of the square-wave driving waveshape now has a maximum possible conversion efficiency of 0.635 (practical efficiency of 0.60 if $V_{sat} = 0.05V_{CC}$) in comparison to a broadband practical efficiency of 0.95. For a case where $P_{dc} = 10$ W the practical device dissipation is $\frac{1}{2}$ W for broadband maximum drive and 4 W for narrowband maximum drive. This eightfold increase in actual dissipation with a maximum amplitude square-wave drive is the significant difference between the broadband and the narrowband cases. The important difference between the sine- and square-wave current pulse drives into a tuned-circuit load is that for the same supply voltage and the same maximum allowable device current the square-wave drive case yields $4/\pi$ times as much output power.

Table 9.1–1

Sine wave	Square wave
$i_d = I_{dc} + I_1 \cos \omega_0 t$	$i_d = I_{dc} +$  $I_p \updownarrow$
$P_{dc} = V_{CC}I_{dc}$	$P_{dc} = V_{CC}I_{dc}$
$I_1 = I_1$	$I_1 = \frac{2}{\pi} I_p;\ I_p = I_{max}$
$V_1 = I_1 R \le V_{CC} - V_{sat}$	$V_1 = \frac{2I_pR}{\pi} \le V_{CC} - V_{sat}$
$P_{ac} = \frac{I_1^2 R}{2}$	$P_{ac} = \frac{2I_p^2 R}{\pi^2}$
$\eta = \frac{1}{2}\left(\frac{I_1}{I_{dc}}\right)^2\left(\frac{V_{CC} - V_{sat}}{V_{CC}}\right)$	$\eta = \frac{1}{2\pi}\left(\frac{I_p}{I_{dc}}\right)^2\left(\frac{V_{CC} - V_{sat}}{V_{CC}}\right)$
$P_{device} = P_{dc}(1 - \eta)$	$P_{device} = P_{dc}(1 - \eta)$

9.2 CLASS B LINEAR RF AMPLIFIERS

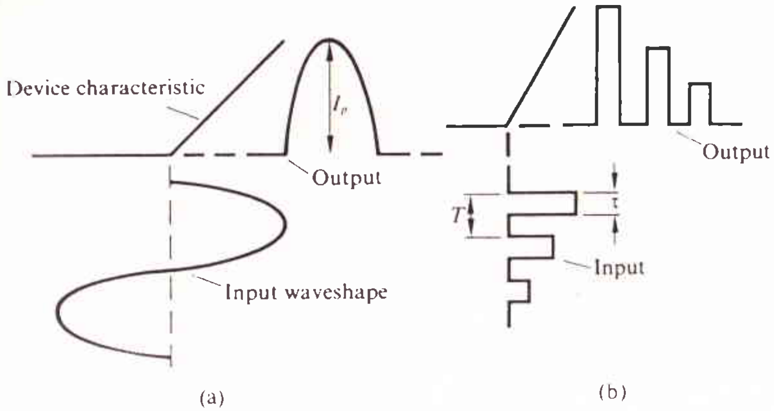
When the signal to be power-amplified is amplitude-modulated in one of the forms discussed in Chapter 8, there must be a linear relationship between the envelope of the power amplifier output and the envelope of the input driving signal.

This desired linearity of the envelope transfer function does not preclude the proper piecewise-linear operation of the device and does not require "undistorted" current or voltage waveshapes throughout the system. What is required is that there be a linear relationship between the amplitude of the fundamental component of the device output current pulses and the input envelope. With such a relationship and with a narrowband output filter, one can obtain a useful power amplifier for SSB, suppressed carrier or normal AM signals.

By using a current pulse train instead of keeping the device current always on, one expects to reduce device dissipation and increase conversion efficiency. This increase in efficiency should occur because current flow will be stopped during the t of the period when the device voltage is high and/or the device voltage will be uced during the part of the cycle when the current is high. Two possible current se shapes are half sine waves and a train of rectangular pulses of width $\tau(\tau < T$ ere $\omega_0 = 2\pi/T)$.

Figure 9.2–1 indicates two idealized possibilities for a combination of device racteristic and driving waveshape that, when followed by an appropriate narrow-nd filter, will allow "linear" RF power amplification.

The fundamental component of the half-sine-wave tips is equal to $I_p/2$; hence,



. 9.2–1 Two possibilities for device characteristics and driving waveshapes that will allow near" RF amplification.

he device characteristic is straight all the way down to the origin, one may "linearly" nplify a 100% modulated AM wave in a single-ended Class B amplifier with a sinsoidal driving waveshape. For any particular value of I_p for which

$$< I_{\max}, \qquad I_p \le \frac{V_{CC} - V_{\text{sat}}}{R_L}, \qquad I_{\text{dc}} = \frac{I_p}{\pi}, \qquad I_1 = \frac{I_p}{2}, \qquad \text{and} \qquad V_1 = I_1 R,$$

that

$$P_{\text{dc}} = \frac{V_{CC} I_p}{\pi}, \qquad P_{\text{ac}} = \frac{I_p^2}{8} R, \qquad \text{and} \qquad \eta = \frac{I_p R_L}{V_{CC}} \frac{\pi}{8},$$

d since $R_L = 2[V_{CC} - V_{\text{sat}}]/I_{\max}$, to achieve optimum efficiency

$$\eta = \frac{I_p}{I_{\max}} \left(\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \right) \frac{\pi}{4}, \tag{9.2–1}$$

ich has a maximum value of $(\pi/4)(V_{CC} - V_{\text{sat}})/V_{CC}$ when $I_p = I_{\max}$. Figure 9.2–2 licates the operating path for this case and compares it with the Class A case own dashed).

It is apparent from this figure why the Class B device dissipation should be reduced with respect to the Class A dissipation. For the high-voltage half of v_d the device carries no current at all. This crowding of the current into the low-voltage half of the cycle is intuitively exactly what one wants in order to reduce the power loss in the device.

It is also apparent that maximum device dissipation will no longer occur at the maximum drive point but at some intermediate point where the product of the device current and voltage drop goes through a maximum. To find this maximum dissipation point, one differentiates $P_{\text{device}} = P_{\text{dc}} - P_{\text{ac}}$ with respect to I_p and sets the result equal to zero:

$$\frac{I_{p_{\text{mdd}}}}{I_{\max}} = \frac{2}{\pi} \frac{V_{CC}}{V_{CC} - V_{\text{sat}}}, \qquad P_{\text{device}_{\max}} = \frac{V_{CC} I_{\max}}{\pi^2} \frac{V_{CC}}{V_{CC} - V_{\text{sat}}}. \tag{9.2–2}$$

Thus for equal values of V_{CC} and $I_{\max}$ the Class B stage has a maximum device dissipation of approximately $\frac{1}{5}$ that of a Class A stage.

Put another way, the ratio of the maximum load power, P_{ac}, to the maximum device dissipation is

$$\begin{aligned} \frac{P_{\text{ac}_{\max}}}{P_{\text{device}_{\max}}} &= \frac{\pi^2}{4} \left(\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \right)^2 && \text{for Class B,} \\ \frac{P_{\text{ac}_{\max}}}{P_{\text{device}_{\max}}} &= \frac{1}{2} \frac{V_{CC} - V_{\text{sat}}}{V_{\text{sat}}} && \text{for Class A,} \end{aligned} \tag{9.2–3}$$

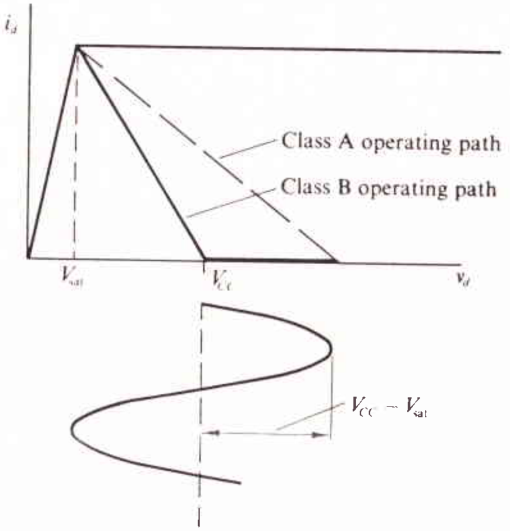


Fig. 9.2–2 Class A and Class B operating paths (sinusoidal input signals).

vhen $V_{sat} = 0.05\,V_{CC}$, the ratio for the sinusoidally driven Class B stage is 2.2 pared to 0.475 for the Class A stage. Thus to supply a maximum load power of W via a Class A stage requires a device capable of dissipating 210 W, while a s B stage requires only a 45 W device.

CLASS C "LINEAR" AMPLIFIERS

h a sinusoidal driving signal, when one attempts to increase the conversion effici- y beyond $\pi/4$ by reducing the conduction angle of the current pulses to less than °, one loses the linear relationship between the fundamental component of the ses and their peak value. The reason is that in this case the conduction angle, and ce the ratio of the fundamental current to the peak current, becomes a function of amplitude.

On the other hand, for a rectangular pulse train of any finite width the fundamen- amplitude is related to the peak amplitude by

$$I_1 = \frac{2I_p}{\pi}\sin\frac{\pi\tau}{T} \tag{9.3–1}$$

hat a linearly modulated pulse AM signal at the input of a zero-biased piecewise- ear stage would yield a linearly modulated AM signal across its output tank circuit.

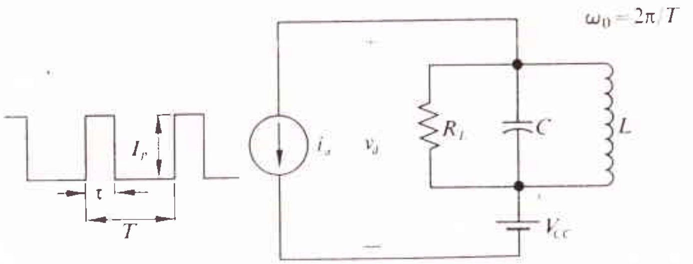


Fig. 9.3–1 RF amplifier with pulse drive.

If one considers the case of a fixed-amplitude pulse train into an output-tuned cuit as shown in Fig. 9.3–1, then, for $I_{1_{\max}}$ $R_L = V_{CC} - V_{sat}$,

$$R_L = \frac{V_{CC} - V_{sat}}{I_{\max}}\frac{\pi}{2}\frac{1}{\sin(\pi\tau/T)},$$

$$P_{ac} = \left(\frac{I_p}{I_{\max}}\right)^2\left(\sin\frac{\pi\tau}{T}\right)\frac{(V_{CC} - V_{sat})I_{\max}}{\pi}, \tag{9.3–2}$$

$$P_{dc} = \frac{I_p}{I_{\max}}\frac{\tau}{T}I_{\max}V_{CC}, \tag{9.3–3}$$

d

$$\eta = \frac{I_p}{I_{\max}}\cdot\frac{V_{CC} - V_{sat}}{V_{CC}}\frac{\sin(\pi\tau/T)}{\pi\tau/T}. \tag{9.3–4}$$

For a given value of $I_p/I_{\max}$ the available output power increases as a sinusoidal function as τ/T increases from zero to $\frac{1}{2}$. At the same time, the circuit conversion efficiency decreases as a $\sin x/x$ function. Table 9.3–1 indicates the relative P_{ac}, efficiency, and in $1 - (\sin x/x)$ the relative device dissipation for a pulse-driven Class C amplifier with the conduction angle as a variable. Note that increasing the conduction angle from 120° to 180° increases the output power by only 15% but increases the device dissipation 1.56 times. From these data one would normally expect to operate with τ/T in the range from $\frac{1}{6}$ to $\frac{1}{3}$ ($120° \geq \theta \geq 60°$).

Table 9.3–1

θ	$\frac{\tau}{T}$	$\sin\frac{\pi\tau}{T}$	$\frac{\sin(\pi\tau/T)}{\pi\tau/T}$	$1 - \frac{\sin x}{x}$
180°	$\frac{1}{2}$	1.00	0.635	0.265
120°	$\frac{1}{3}$	0.866	0.83	0.17
90°	$\frac{1}{4}$	0.707	0.90	0.10
60°	$\frac{1}{6}$	0.500	0.955	0.045

Note that these data are perfectly valid if the bias point is chosen anywhere to the left of the breakpoint in Fig. 9.2–1(b). The only time that breakpoint bias operation is necessary is when "linear" envelope amplification is required.

As in the Class B case of sine-wave drive with a constant supply voltage, when the device operates in an on-off fashion the maximum device dissipation does not occur at the maximum drive but at some intermediate level. Again this maximum level is found by writing $P_{device} = P_{dc} - P_{ac}$, differentiating with respect to I_p, and setting the result equal to zero. When this is done for the case of rectangular pulses the point of maximum device dissipation, $I_{p_{mdd}}$, is found as

$$\frac{I_{p_{mdd}}}{I_{\max}} = \frac{1}{2}\frac{V_{CC}}{V_{CC} - V_{sat}}\frac{\sin(\pi\tau/T)}{\pi\tau/T}. \tag{9.3–5}$$

At this drive level the actual efficiency is

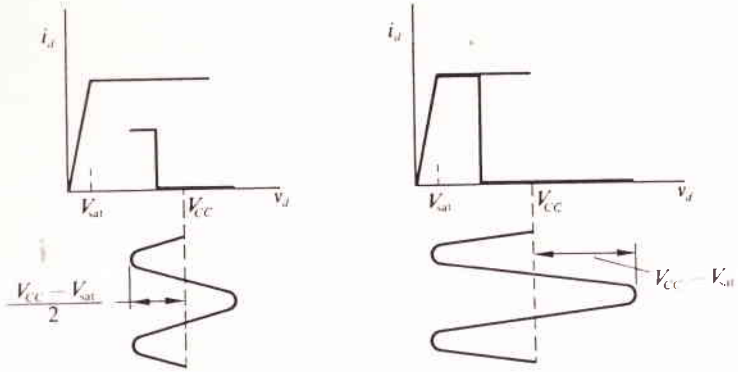
$$\eta = \frac{1}{2}\left[\frac{\sin(\pi\tau/T)}{\pi\tau/T}\right]^2, \tag{9.3–6}$$

which for $\theta = 90°$ and $V_{sat} = V_{CC}/20$ is down from 0.855 at maximum drive to only 0.405 of the maximum device dissipation point of $I_{p_{mdd}}/I_{\max} = 0.475$.

Figure 9.3–2 indicates the operating paths for no modulation and for 100% positive peak modulation for a 90° rectangular pulse drive Class C "linear" amplifier.

It will be demonstrated in Section 9.9 that the overall efficiency can be increased if the output voltage is varied in step with the input. The general results of that section can also be applied to the pulse-driven linear Class C amplifiers of this section.

For ordinary power amplification of unmodulated signals, Class C is preferred over Class B. At higher frequencies the pulse-like drive signals are difficult to produce



No modulation Modulation peak

Fig. 9.3–2 Operating paths for 90° "linear" Class C amplifier.

d one tends to use sinusoidally driven Class C amplifiers. Such amplifiers are scussed in Section 9.4.

With a triode vacuum tube the output current pulse is a function of the output oltage; hence direct output modulation of such stages is possible. Furthermore, ey can be adjusted for reasonably linear operation up to nearly 100% modulation. herefore, vacuum tube amplifiers are modulated in the final power amplifier whenver possible. Normally single-sideband and suppressed carrier operations are perrmed at low power levels. These signals do require linear power amplifiers for all bsequent stages.

Transistors would seem to lend themselves quite readily to the modulated Class amplifier scheme of Section 9.9. Since the best-modulated Class C amplifier eldom has an overall efficiency of more than 80%, one is really sacrificing almost othing in this case.

Of course, FM, PM, binary AM, and the various industrial uses of CW are comletely indifferent to the "linearity" of the power amplifier employed.

.4 RF CLASS C AMPLIFIERS

The normal RF power amplifier does not have a piecewise-linear device characteristic, or is it driven by a "square" pulse. It is normally operated into a tuned-circuit load, o that no matter what the shape of the input current pulse one can begin by assuming hat the output voltage is sinusoidal. If one knows the actual input driving waveshape, f one knows that the output voltage is sinusoidal, and if one has the proper device haracteristic, then one can always calculate the output current waveshape.

In some cases this current pulse will have a known analytic form, and the dc and fundamental terms may be written by inspection. Even if the analytic form is not known, one can normally approximate any symmetrical pulse by one of several forms such as rectangular, trapezoidal, triangular, cosine, cosine squared, or sine-wave tip. (If the output circuit is tuned to be resistive, then a symmetrical input pulse will yield a symmetrical output pulse.)

To aid in the analysis of the current pulses, the appendix to this chapter tabulates the equations for various pulse forms and presents normalized curves for I_{dc}, I_1, and I_2 vs. the pulse width or conduction angle. As these curves indicate, the variations among the differential waveshapes are not large; consequently an exact fit is not necessary in order to obtain a reasonable estimate of the fundamental and dc values.

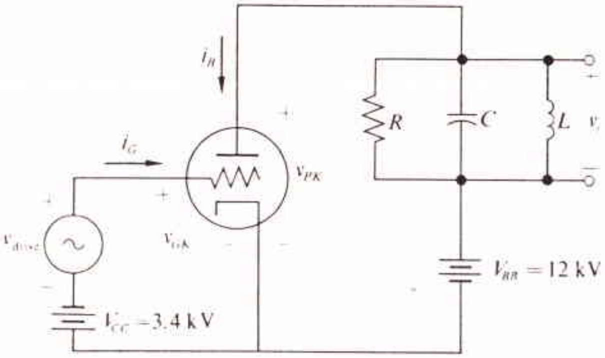


Fig. 9.4–1 Triode Class C amplifier.

The circuit of Fig. 9.4–1 provides an example of the use of these curves as well as of the determination of the current waveshape for an idealized triode vacuum tube in which the plate current is a function of both v_{PK} and v_{GK}. In order to simplify our example slightly, let us assume that the device characteristics are as shown in Fig. 9.4–2. (As shown in Fig. 9.4–2, this tube has a constant $\mu = 10$ and a constant $r_p = 1.8\ \text{k}\Omega$. In a real tube the curves would be neither straight nor equally spaced.) The bias has been chosen so that with a sine-wave drive of 5.4 kV peak and with a $V_{\min} = +2$ kV for the plate-cathode voltage, the conduction angle will be 120°.†

For comparison purposes let us assume three different driving waveshapes. In each case the bias and supply voltages are constant but R_L is changed, so that we always drive up to $v_{GK} = v_{PK} = 2$ kV in the upper left-hand corner:

1. Sinusoidal: 5.4 kV, $\theta_p = 120°$
2. Pulse: 5.4 kV, $\theta_p = 120°$
3. Composite: (5.4 kV) $(1.1 \cos \omega t - 0.1 \cos 3\omega t)$, $\theta_p \approx 140°$

The operating paths for the three cases are shown in Fig. 9.4–3. These paths are constructed by drawing the driving voltage along the vertical axis centered at -3.4 kV,

---

† For idealized tubes the relation among V_{BB}, V_{CC}, μ, and the conduction angle θ_p may be written as

$$V_{CC} = -\left[\frac{V_{BB}}{\mu} + \left(\frac{v_{PK\min} + \mu v_{GK\max}}{\mu}\right)\frac{\cos(\theta_p/2)}{1 - \cos(\theta_p/2)}\right]; \tag{9.4–1}$$

hence for $V_{BB} = 12$ kV, $\mu = 10$, $v_{PK_{\min}} = 2$ kV, $v_{GK_{\max}} = 2$ kV, and $\theta_p = 120°$, it follows that $V_{CC} = -3400$ V.

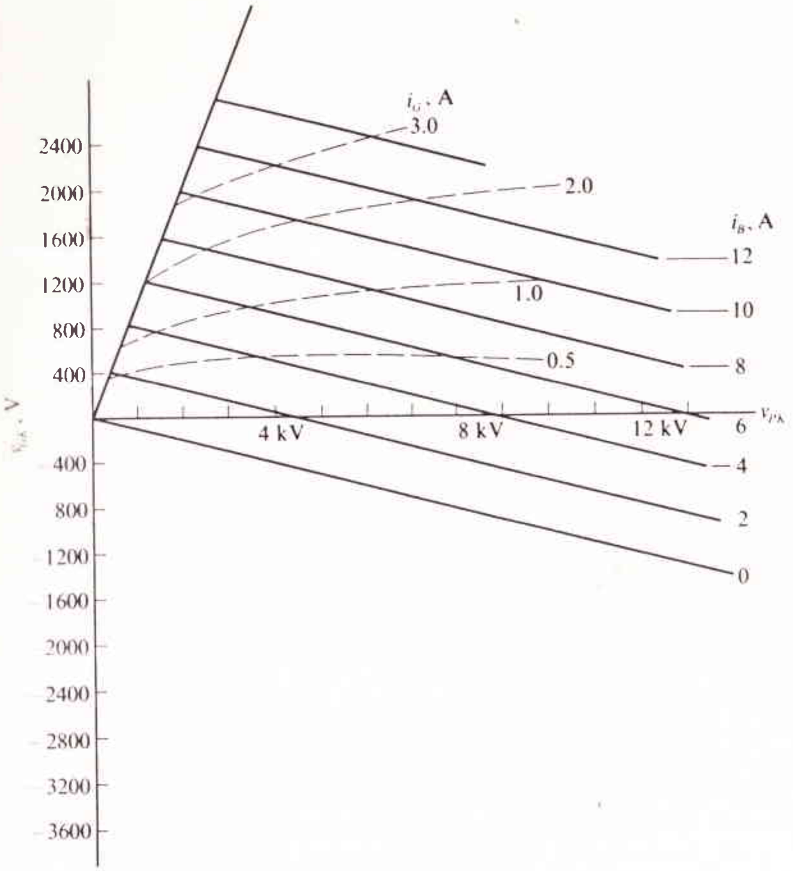


Fig. 9.4–2 Mythical triode characteristics ($\mu = 10$, $r_p = 1800\ \Omega$).

wing a 10 kV sine wave along the horizontal axis centered around 12 kV, and tting the resultant Lissajous pattern. All the circled points in any vertical line resent the same angular distance measured from the peak of the current pulse from the valley of the output tank circuit voltage). Note, as in previous cases, the reme dependence of the operating path on the driving waveshape.

Once the operating path is sketched, the current pulse is drawn almost by pection, since each circle indicates the plate or grid current for its particular gular displacement. The respective current pulses are shown in Fig. 9.4–4.

In order to compare the three cases let us approximate the various current pulses follows.

Sinusoidal drive:

i_B—cosine pulse of 120° width, $I_1 = 3.8$ A, $I_{dc} = 2.1$ A;

i_G—triangular pulse of 102° width, $I_{G1} = 0.85$ A, $I_{G0} = 0.45$ A.

2. Pulse drive:

   i_B—square pulse of 120° width minus a triangular pulse of 120° width and 0.2 A height, $I_1 = 5.5 - 0.6 = 4.9$ A, $I_{dc} = 3.33 - 0.33 = 2.0$ A;

   i_G—square pulse of 120° width and 2.2 A height plus a triangular pulse of 120° width and 1 A height, $I_{G1} = 1.2 + 0.30 = 1.5$ A, $I_{G0} = 0.733 + 0.167 = 0.90$.

3. Composite drive:

   i_B—symmetrical trapezoid of 140° width, $I_1 = 5$ A, $I_{dc} = 2.0$ A;

   i_G—cosine pulse of 120° width, $I_{G1} = 1.14$ A, $I_{G3} = 0.42$ A, $I_{G0} = 0.63$ A.

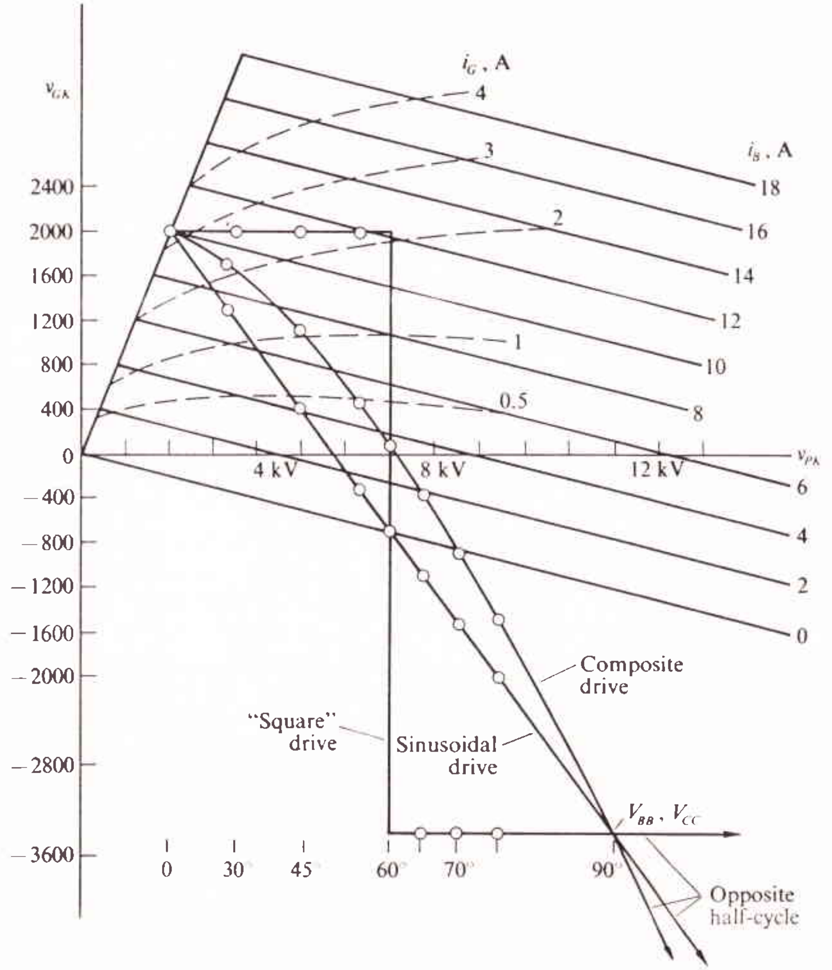


Fig. 9.4–3 Operating paths for three different driving waveshapes and load resistors ($V_{CC} = 3400$ V, $V_{BB} = 12$ kV).

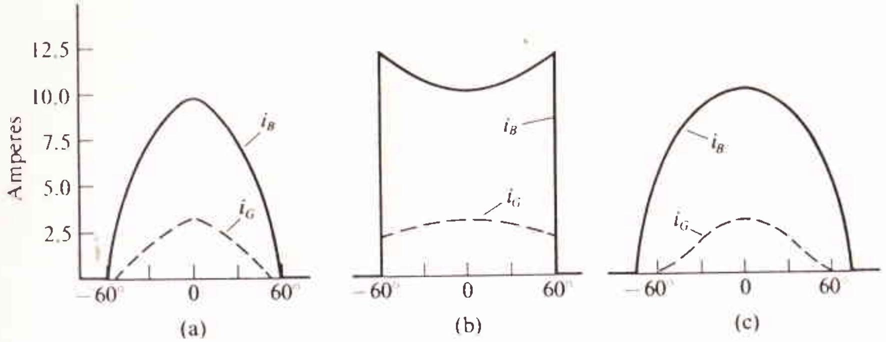


Fig. 9.4–4 Plate and grid current pulses for the three operating paths of Fig. 9.4–3. (a) Sinusoidal drive. (b) "Pulse" drive. (c) Composite drive.

From these current waveshapes and their harmonic analyses in addition to the assumption of a known driving waveshape and a known output waveshape (10 kV peak sinusoid), we can derive values of i_B and i_G for the necessary value of R_L as well as the various powers listed in Tables 9.4–1 and 9.4–2.

The grid circuit calculations are explained in Section 9.8.

Which driving waveshape is allowable will depend on the device limitations. For example, if the peak allowable current† is 10 A, then the pulse drive would not be satisfactory although the other two would be. (Actually the peak cathode currents do not differ too drastically, since for the pulse case the valley in i_B corresponds to the peak in i_G.) If the peak plate dissipation is 7 kW or the peak grid dissipation is 900 W, then only the sinusoidal case would be allowed. The composite drive does yield the same output power as the "pulse" drive, but does it at a higher efficiency, with 20% less peak current, with 76% of the grid dissipation, and with a waveshape that is much easier to generate than is the pulse waveshape. The last column of Table 9.4–2 indicates the power gain of the stage. From this and from the driver power figure it is apparent that a power amplifier will be needed to drive this stage.

Table 9.4–1 Plate circuit values

Driving waveshape	R_L, Ω	P_{ac}, kW	P_{dc}, kW	P_{plate}, kW	η
Sinusoidal	2620	19	25.2	6.2	0.75
Pulse	2040	24.5	36	11.5	0.685
Composite	2000	25	35.5	10.5	0.705

Table 9.4–2 Grid circuit values

Driving waveshape	$P_{\text{from driver}}$, W	$P_{\text{to battery or bias circuit}}$, W	P_{grid}, W	P_{ac}/P_{driver}
Sinusoidal	2300	1530	770	8.3
Pulse	4850	3050	1800	5.05
Composite	3400	2253	1147	7.1

It is apparent that one could generate a series of curves of P_{ac}, η, and P_{device} vs. conduction angle for different assumed device current waveshapes. In general, the results will all be similar to those of Table 9.3–1 for the rectangular pulse. That is, the power output will rise as the conduction angle increases, but the device dissipation will rise even more rapidly; hence normally the conduction angle must be held below 180°.

In order to increase the efficiency while maintaining the power output, one wants to operate the device completely as a switch so that the device voltage drop remains at or near the minimum value over a half-cycle of current flow. The next section examines such switched-device or Class D circuits.

9.5 NARROWBAND CLASS D POWER AMPLIFIERS

The two basic forms of the RF Class D power amplifier are shown in Fig. 9.5–1. In both cases the switch is driven back and forth at the resonant frequency of the tuned circuit. Initially the time when neither contact is made is assumed to be zero.

Voltage-Switching Case—Ideal (100% Efficiency)

Since the switch spends half its time in each position, the voltage $v_i(t)$ is a square wave of amplitude $2V_{CC}$. This square wave can be expanded by means of Eq. (4.2–4) to obtain a dc value of V_{CC} and a fundamental value of $V_1 = 4V_{CC}/\pi$. Since the circuit is resonant at the fundamental frequency of the square wave and is assumed to have $Q_T \geq 5$, the current is essentially a sinusoid of peak amplitude $I_1 = V_1/R$. Since this current must also flow through the switches, it follows that the current through each

---

† The maximum cathode current should be supplied by the manufacturer of the device. If it is not, then the most reasonable procedure is to find the peak allowable cathode current in milliamperes per watt of filament power for another tube having the same type of filament material and roughly the same power rating. This "constant" can then be used to determine $I_{K_{max}}$ for the tube in question.

For pure tungsten filaments run at their rated power input, this value is usually about 5 mA/watt. As the filament power decreases, the emission per watt falls but the life increases. Thus reducing the filament power to 90% of its rated value (normally this means reducing the filament voltage to about 95% of its rated value) will approximately double the life of the filament while reducing the maximum emission per watt to perhaps 75% of its original value.

The lives of thoriated tungsten filaments cannot be extended by low-power operation. Normally the filament voltage for these tubes should be regulated to within 5% of the nominal value. Such filaments are much more efficient emitters than pure tungsten. However, they should not be operated within a factor of 2 or 3 of their temperature saturation currents; hence one has a nominal mA/watt figure in the region of 15 to 17 for this type of filament.

Oxide-coated cathodes are generally too delicate to withstand the ionic bombardments of high-voltage operation. With small tubes that have oxide-coated cathodes one can assume that peak cathode currents of 40 mA/watt will not cause damage.

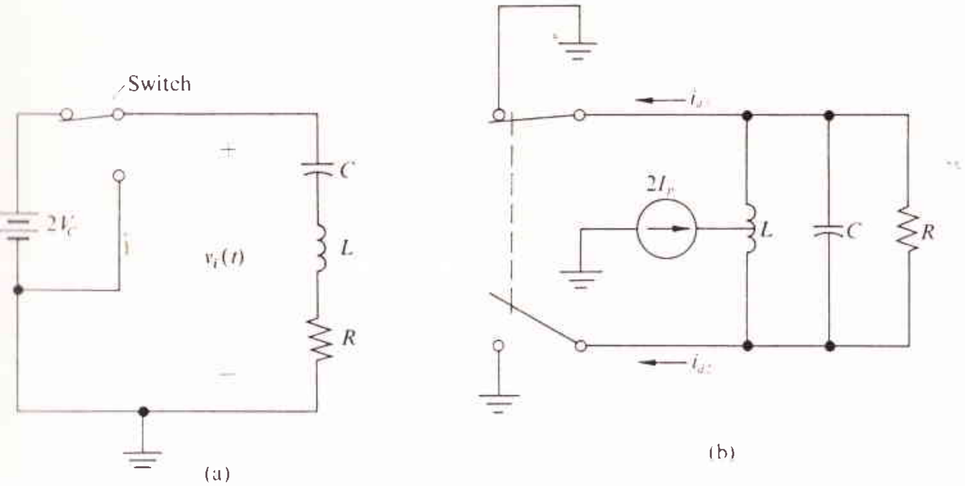


Fig. 9.5–1 (a) Voltage-switching Class D circuit. (b) Constant-current-switching Class D circuit.

witch is a half sine wave. However, once the current through the battery is known, hen P_{dc} can be calculated and the analysis problem is solved.

oltage-Switching Case—Practical

he efficiency of this circuit departs from 100% because of losses in the switches and the inductor and/or capacitor. Since the device is assumed always to be saturated hen on, the closed switches may be modeled as shown in Fig. 9.5–2.

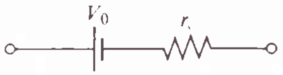


Fig. 9.5–2 Device model in saturated region.

Now if r_L is the combined effective series loss element of the inductor and the apacitor and if $R_\Sigma = r_s + r_L + R$, then the effective peak-to-peak input voltage quare wave is reduced in magnitude to $2(V_{CC} - V_0)$, from which

$$V_1 = \frac{4(V_{CC} - V_0)}{\pi}, \qquad I_1 = \frac{V_1}{R_\Sigma}, \qquad \text{and} \qquad I_{dc} = \frac{I_1}{\pi}.$$

bviously the devices must not break down at a voltage of $2V_{CC}$; on the other hand, they are to be used to the limits of their current, then the peak current I_1 will be set at $_{max}$ (certainly $I_1 \leq I_{max}$).

With these assumptions.

$$P_{dc} = 2V_{CC}I_{dc}, \qquad P_{ac} = \frac{V_1 I_1}{2} = \frac{4(V_{CC} - V_0)}{\pi}\frac{I_{dc}\pi}{2},$$
$$\eta_{device} = \frac{P_{ac}}{P_{dc}} = \frac{V_{CC} - V_0}{V_{CC}}, \tag{9.5–1}$$

---

and

$$\eta_{overall} = \frac{R}{R_\Sigma}\frac{V_{CC} - V_0}{V_{CC}}. \tag{9.5–2}$$

If $I_1 = I_{max}$, then

$$R_{\Sigma max} = \frac{4(V_{CC} - V_0)}{\pi I_{max}},$$
$$P_{dc} = \frac{2V_{CC}I_{max}}{\pi}, \qquad P_{ac\,load} = \frac{2(V_{CC} - V_0)}{\pi}I_{max}\frac{R}{R_\Sigma}, \tag{9.5–3}$$

and

$$P_{per\,device} = \frac{V_0 I_{max}}{\pi} + \frac{I_{max}^2 r_s}{4}. \tag{9.5–4}$$

For example, if $I_{max} = 2$ A, $V_{CC} = 20$ V, $V_0 = 0.4$ V, $r_s = 0.25\ \Omega$, $r_L = 0.125\ \Omega$, and $Q_L = 100$, then $R_{\Sigma max} = 12.12\ \Omega$ and the loaded Q of the series circuit is approximately 10. Thus

$$P_{dc} = 25.4 \text{ W}, \qquad P_{ac\,load} = 24 \text{ W}, \qquad \eta = 0.95,$$

and

$$P_{device} = 0.255 + 0.25 = 0.55 \text{ W}.$$

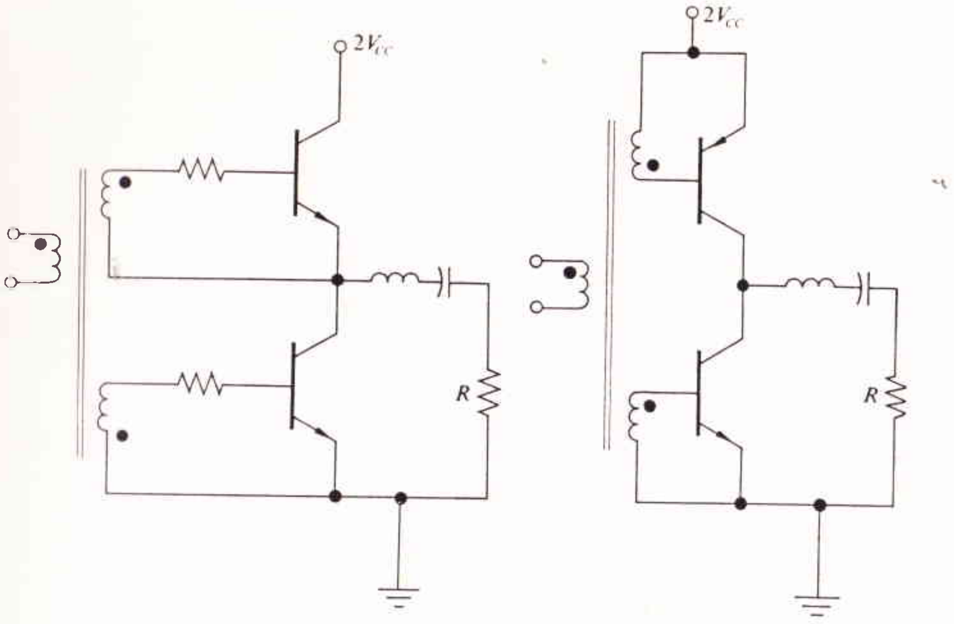
Hence the ratio of ac load power to total device dissipation (two devices) is nearly 22/1; this value should be compared to the ratio of 2.2/1 for the Class B sinusoidally driven stage.

Figure 9.5–3 illustrates two transistor arrangements that would be possible as switches.

As we shall see in a subsequent section, it is a straightforward matter to build a transformerless driver for the complementary symmetry version of the switch. The driving circuit must supply enough turn-on current to keep the on transistor saturated during its conducting half-cycle, and then enough turn-off current to get it out of saturation rapidly. If the base current is sinusoidal and β is constant as the current varies, then $I_{B1} \geq I_1/\beta$, so that if $I_1 = 2$ A and $\beta \approx 10$, then $I_{B1} \geq 200$ mA. The turn-off problem is important because if one unit stays saturated while the second unit comes on, the result is almost a dead short from $2V_{CC}$ to ground. Unless the power supply has a current-limiting feature, the transistors may be destroyed by these repeated surges.

If a square-wave drive is employed, then a "speed-up" capacitor might be added in parallel with the base resistors to increase the turn-off and turn-on times.

As the carrier frequency increases, the assumption of negligible transition times for the switch will no longer be valid. Since a heavy current flows during this interval while the switch is going off, the rise time does not have to become appreciable before the transition-region losses are equal to the saturation-region terms and the very



(a) (b)

Fig. 9.5-3 Two possible voltage-switching circuits.

ow calculated values for P_{device} are doubled or tripled. In estimating the transition osses, we can assume linear rise and fall times, each of duration t_t. Then the total ansition-region power loss will be approximately $(t_t/T)P_{\text{dc}}$. Thus if the device ower for our previous example is not to be more than doubled, we need $t_t/T \leq 0.02$. his means that for operation at 500 kHz, $t_t \leq 40$ nsec, and for operation at 1 MHz, ≤ 20 nsec.

Since these numbers represent approximately the limits of present-day power ansistors, one can hope to achieve ratios of load power to total device loss of about 0/1 only up to center frequencies of about 1 MHz. For lower frequencies the gains ver the Class B or Class C stage can be impressive.

urrent-Switching Case—Ideal

or the circuit of Fig. 9.5–1(b) the current through each of the two switches has the rm of a square wave of amplitude $2I_p$. These square waves may be expanded to ovide a dc value of I_p through each switch and a fundamental component $4I_p/\pi$ hich, because of the phase reversal between the two switch currents, may be viewed flowing out of the bottom of the tank circuit and into the top. That is, from the ewpoint of currents, the circuit of Fig. 9.5–1(b) may be replaced by that of Fig. 5–4.

Since the tuned circuit is presumed to have a reasonable Q, we neglect the harmonic generators; and since the switch is presumed to spend equal times on each

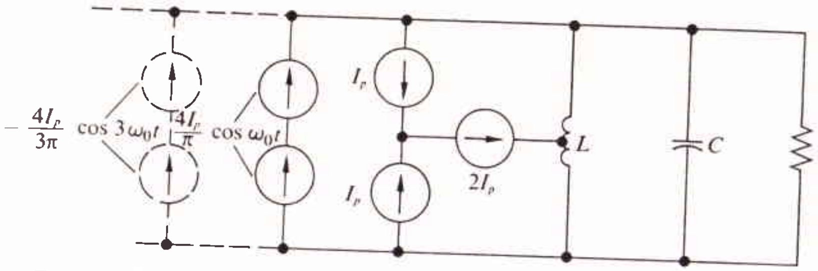

$-\frac{4I_p}{3\pi}\cos 3\omega_0 t$
$\frac{4I_p}{\pi}\cos \omega_0 t$
I_p
I_p
$2I_p$
L
C
R

Fig. 9.5–4 Equivalent circuit (from the current viewpoint) for Fig. 9.5–1 (b).

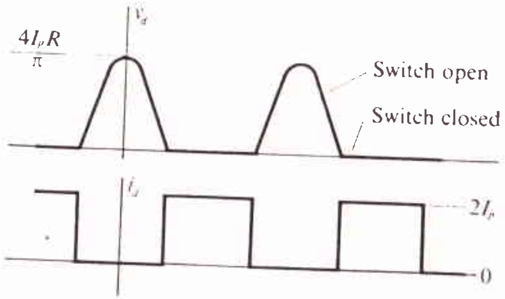


Fig. 9.5–5 Device current and voltage in the current-switching circuit.

half, we combine the fundamental generators. Now $V_1 = 4I_pR/\pi$ and the tank circuit voltage is sinusoidal. The voltage across each device is not sinusoidal but takes the form of a train of half-sine-wave pulses as shown in Fig. 9.5–5.

Current-Switching Case—Practical

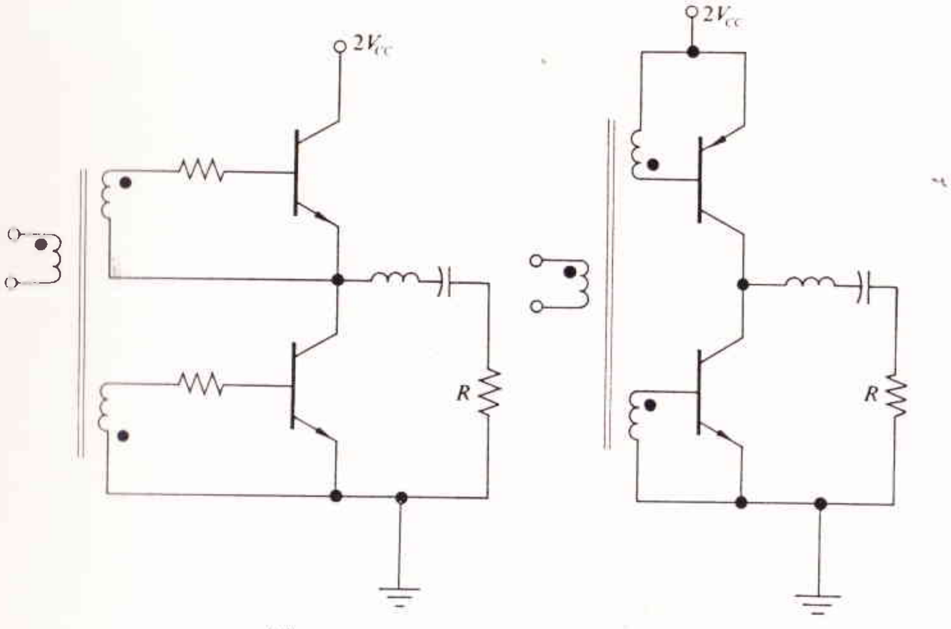
In a real circuit the constant current supply must be able to adjust itself to the needs of the circuit; hence it would be mechanized by an inductance, L_{CC}, that is at least five times as great as L in series with a supply voltage V_{CC}. Then, since $P_{\text{dc}} = 2I_pV_{CC}$ while the ac power removed is $8I_p^2R/\pi^2$, it follows that, for the ideal case where there is no device loss, $P_{\text{dc}} = P_{\text{ac}}$ and

$$2I_p = \frac{\pi^2 V_{CC}}{2R}. \tag{9.5–5}$$

If we wish to set $2I_p$ at $I_{\max}$ for each switch, then we must set R at a particular value. Once R is set, all the powers and efficiencies can be calculated immediately.

In this case the current during conduction is constant; hence for any particular $I_{\max}$ one can simplify the model of Fig. 9.5–2 back to a single battery of value V_{sat}. Again the L and C losses could be represented as an equivalent loss R_l; thus the tank circuit efficiency is $R_l/(R_l + R)$ and the power into the tank is

$$\left(\frac{4I_p}{\pi}\right)^2 \frac{R_{\parallel}}{2},$$



(a) (b)

Fig. 9.5-3 Two possible voltage-switching circuits.

w calculated values for P_{device} are doubled or tripled. In estimating the transition sses, we can assume linear rise and fall times, each of duration t_t. Then the total ansition-region power loss will be approximately $(t_t/T)P_{dc}$. Thus if the device ower for our previous example is not to be more than doubled, we need $t_t/T \leq 0.02$. his means that for operation at 500 kHz, $t_t \leq 40$ nsec, and for operation at 1 MHz, ≤ 20 nsec.

Since these numbers represent approximately the limits of present-day power ansistors, one can hope to achieve ratios of load power to total device loss of about)/1 only up to center frequencies of about 1 MHz. For lower frequencies the gains ver the Class B or Class C stage can be impressive.

urrent-Switching Case—Ideal

or the circuit of Fig. 9.5–1(b) the current through each of the two switches has the rm of a square wave of amplitude $2I_p$. These square waves may be expanded to ovide a dc value of I_p through each switch and a fundamental component $4I_p/\pi$ hich, because of the phase reversal between the two switch currents, may be viewed flowing out of the bottom of the tank circuit and into the top. That is, from the ewpoint of currents, the circuit of Fig. 9.5–1(b) may be replaced by that of Fig. 5–4.

Since the tuned circuit is presumed to have a reasonable Q, we neglect the harmonic generators; and since the switch is presumed to spend equal times on each half, we combine the fundamental generators. Now $V_1 = 4I_pR/\pi$ and the tank circuit voltage is sinusoidal. The voltage across each device is not sinusoidal but takes the form of a train of half-sine-wave pulses as shown in Fig. 9.5–5.

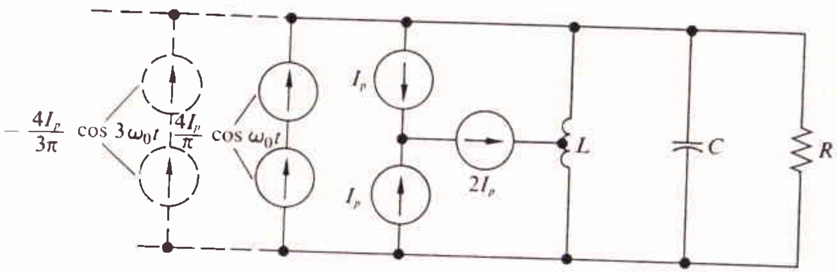

$-\frac{4I_p}{3\pi}\cos 3\omega_0 t$
$\frac{4I_p}{\pi}\cos \omega_0 t$
I_p
I_p
$2I_p$
L
C
R

Fig. 9.5–4 Equivalent circuit (from the current viewpoint) for Fig. 9.5–1 (b).

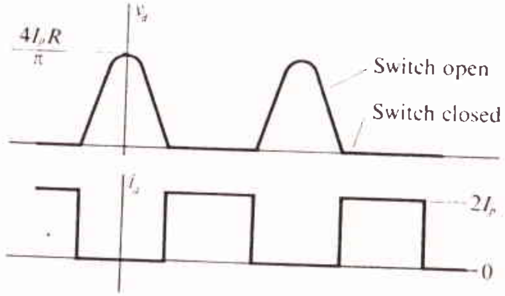


Fig. 9.5–5 Device current and voltage in the current-switching circuit.

Current-Switching Case—Practical

In a real circuit the constant current supply must be able to adjust itself to the needs of the circuit; hence it would be mechanized by an inductance, L_{CC}, that is at least five times as great as L in series with a supply voltage V_{CC}. Then, since $P_{dc} = 2I_pV_{CC}$ while the ac power removed is $8I_p^2R/\pi^2$, it follows that, for the ideal case where there is no device loss, $P_{dc} = P_{ac}$ and

$$2I_p = \frac{\pi^2 V_{CC}}{2R}. \tag{9.5–5}$$

If we wish to set $2I_p$ at $I_{\max}$ for each switch, then we must set R at a particular value. Once R is set, all the powers and efficiencies can be calculated immediately.

In this case the current during conduction is constant; hence for any particular $I_{\max}$ one can simplify the model of Fig. 9.5–2 back to a single battery of value V_{sat}. Again the L and C losses could be represented as an equivalent loss R_l; thus the tank circuit efficiency is $R_l/(R_l + R)$ and the power into the tank is

$$\left(\frac{4I_p}{\pi}\right)^2 \frac{R_{\parallel}}{2},$$

'e

$$R_{\parallel} = RR_l/(R + R_l) = \eta_{\text{tank}} R,$$

$$P_{\text{dc}} = 2I_p V_{CC}, \qquad P_{\text{ac load}} = \left(\frac{4I_p}{\pi}\right)^2 \frac{\eta_{\text{tank}}^2 R}{2}, \tag{9.5–6}$$

$$P_{\text{each device}} = I_p V_{\text{sat}}, \tag{9.5–7}$$

R is chosen to limit the value of $2I_p$ to a value that the device can handle:

$$R_{\parallel} \geq \frac{\pi^2(V_{CC} - V_{\text{sat}})}{2I_{\max}}. \tag{9.5–8}$$

sing $R_{\parallel}$ above this value will reduce the power available to the load.

In order to compare the voltage- and current-switching cases we might calculate erical values for a set of parameters similar to those used in the voltage-switching . Such values are shown in Table 9.5–1.

Table 9.5–1

Voltage switch	Current switch
$V_{CC} = 20\text{ V}, I_{\max} = 2\text{ A}$	$V_{CC} = 20\text{ V}, I_{\max} = 2\text{ A}$
$V_0 = 0.4\text{ V}, r_s = 0.25\,\Omega$	$V_{\text{sat}} = 0.9\text{ V}$
$r_L = 0.125\text{ k}\,\Omega, Q_L \approx 100$	$R_l = 4.9\text{ k}\Omega, Q_l \approx 100$
Loaded $Q \approx 10$	*Loaded* $Q \approx 10$
$R = 12.1\,\Omega$	$R = 48.5\,\Omega$
$P_{\text{dc}} = 25.4\text{ W}$	$P_{\text{dc}} = 40\text{ W}$
$P_{\text{load}} = 24\text{ W}$	$P_{\text{load}} = 37.8\text{ W}$
$P_{\text{per device}} = 0.55\text{ W}$	$P_{\text{per device}} = 0.9\text{ W}$
$\eta_{\text{overall}} = 0.95$	$\eta_{\text{overall}} \approx 0.95$

The current switch is not subject to the problem of excess current when both itches are closed simultaneously, but it does suffer from excess device voltage if th switches try to open at once. It also suffers from the practical difficulty that the n-off switching starts from the full on current rather than from the theoretical lue of zero device current that should be present in the voltage switch at the instant switching.

If the peak current of the devices is limited, then the square current pulse of the rrent switch will provide more load power than the half sine pulse of the voltage-itching case. On the other hand, if the circuit is device dissipation limited, then in eory the circuits can yield exactly equal outputs. In practice, the actual efficiency the current-switching circuit tends to be lower than the efficiency of the voltage-itching circuit; consequently, if the device dissipation is limited, the voltage switch the better choice.

Other factors to consider when making a choice between the circuits have to do with the potential device failure mechanisms and whether excess current or excess voltage seems more dangerous. More details on breakdown and losses in this type of circuit can be found in a paper by Chudobiak and Page.†

9.6 BROADBAND CLASS B AMPLIFIERS

To supply the baseband power necessary to drive a modulated Class B or Class C stage or to drive the loudspeakers of a public address system or to drive the loudspeaker of a radio or television set, one would like a broadband, high-gain, distortionless amplifier built of the least expensive possible devices.

In Section 9.1 we investigated Class A amplifiers and found that, though their distortion could be low, their overall average efficiency was very low and, in general, their required device dissipation was more than twice their peak ac load power.

In Section 9.2 we combined a device having half-sine-wave output pulses of current with a filter to obtain a "linear" amplifier for the "envelope" of narrowband signals. With broadband signals we cannot use a filter; however, we can combine two devices so that they alternately force half-sine-wave pulses through the load in opposite directions. Such a scheme will give us the advantages of Class B over wide-frequency ranges. Figure 9.6–1 illustrates one of the basic forms of a Class B broadband amplifier.

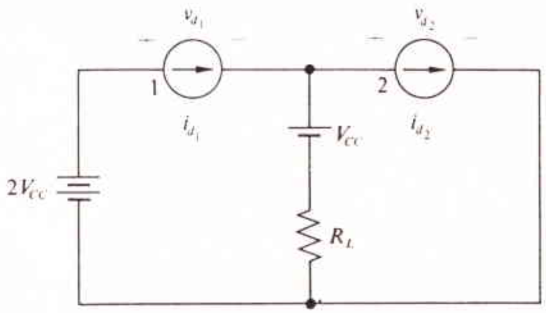


Fig. 9.6–1 Class B amplifier configuration.

If the current waveshapes of generators 1 and 2 are as shown in Fig. 9.6–2, where

$$I_{\max} = \frac{V_{CC} - V_{\text{sat}}}{R_L},$$

is the maximum value for I_{p1} and I_{p2}, then for this maximum drive

$$P_{\text{ac}} = \frac{(V_{CC} - V_{\text{sat}})I_{\max}}{2}, \tag{9.6–1}$$

---

† W. J. Chudobiak and D. F. Page, "Frequency and Power Limitations of Class D Transistor Amplifiers," *IEEE Journal of Solid State Circuits*, SC-4, pp. 25–37 (Feb. 1969).

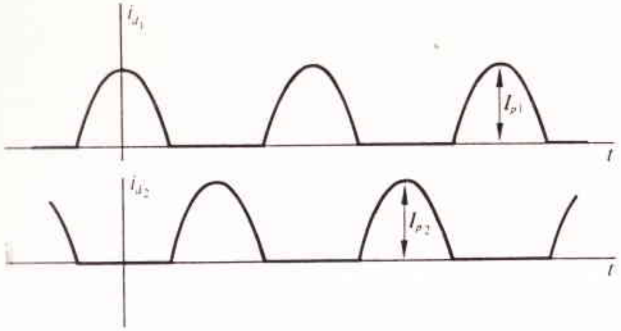


Fig. 9.6–2 Current waveshapes for Class B amplifier with sinusoidal drive.

$$P_{\mathrm{dc}} = 2V_{CC}\frac{I_{\max}}{\pi}, \qquad \eta = \frac{\pi}{4}\frac{V_{CC} - V_{\mathrm{sat}}}{V_{CC}}. \tag{9.6–2}$$

he maximum drive efficiency in this case is 50 % higher than it was in the case of A sinusoidal drive. As was true for narrowband operation, such a comparison sufficient to give a good indication of the relative device dissipation possibilities two cases. As before, the peak device dissipation does not occur at peak drive or a sinusoidal signal, occurs when

$$\frac{I_{P_{\mathrm{mdd}}}}{I_{\max}} = \frac{2}{\pi}\frac{V_{CC}}{V_{CC} - V_{\mathrm{sat}}}, \qquad P_{\mathrm{per\ device_{max}}} = \frac{V_{CC}I_{\max}}{2\pi^2}\frac{V_{CC}}{V_{CC} - V_{\mathrm{sat}}}. \tag{9.6–3}$$

gain we find that for Class B the device (there are two devices now) must dissipate about $\frac{1}{5}$ as much power in the worst case as does a Class A stage that provides me useful peak power to the load. With variable signals such as speech or music. atio of necessary device dissipation is more nearly 20/1 between Class B and A stages.

ecause of this enormous differential in device dissipation, almost no Class A s are used as power amplifiers. Even in small transistor radios where the peak ut power may be only of the order of 50 mW, the low average dissipation and verage power supply current dictate the use of Class B stages.

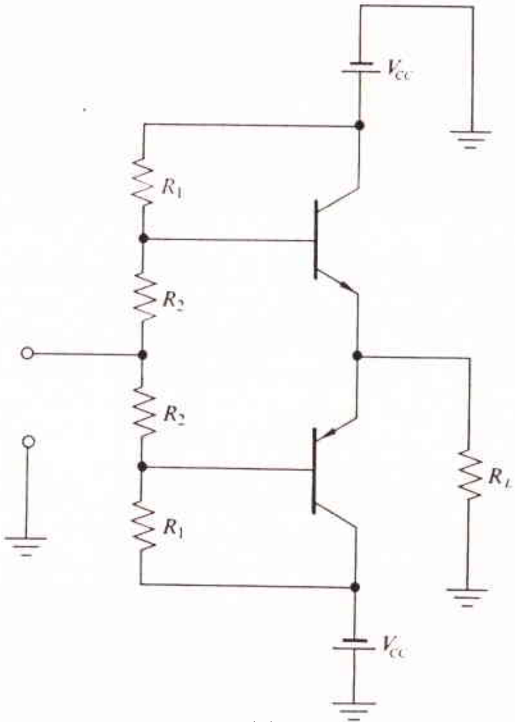
ne basic difficulty with the current-source Class B stages is source matching. xample, we may look on the load current as being made up of two trains of half waves as shown in Fig. 9.3–2. Each of these trains has a dc and an infinity of even onic values, as well as a fundamental value of $I_p/2$. If $I_{p1} = I_{p2}$, then the dc all the harmonics except the fundamental cancel each other and we obtain a sine-wave output. If $I_{p1} \neq I_{p2}$, then this cancellation fails and distortion results. rom Eq. (4.2–1) the resultant second- and fourth-harmonic distortions are seen

$$\frac{\text{cond harmonic}}{\text{fundamental}} = \frac{4}{3\pi}\frac{I_{p1} - I_{p2}}{I_{p1} + I_{p2}}, \qquad \frac{\text{fourth harmonic}}{\text{fundamental}} = \frac{4}{15\pi}\frac{I_{p1} - I_{p2}}{I_{p1} + I_{p2}}. \tag{9.6–4}$$

Thus if the current sources are transistors with a 2/1 mismatch of β's (not at all unlikely with unselected units), one will have 14 % second-harmonic distortion and 2.9 % fourth-harmonic distortion from this cause alone. To keep the second-harmonic distortion term from mismatch to below 1 %, the sources should be matched to within 2.5 %.

Many smaller (less than 1 W) transistors from the same batch (certainly from the same chip) will meet this matching criterion without undue selection.

Figure 9.6–3 shows four possible broadband Class B stages. The two emitter-follower stages avoid the β-matching problem at the expense of providing only unity voltage gain. (We assume that the stage driving the stages shown has an output impedance that is small in comparison to βR_L. Another emitter follower can usually provide such a driven stage.) If reasonably symmetrical PNP and NPN transistors can be found,† then, at the expense of two separate power supplies, the circuit of Fig. 9.6–3(a) avoids all transformers and coupling capacitors.

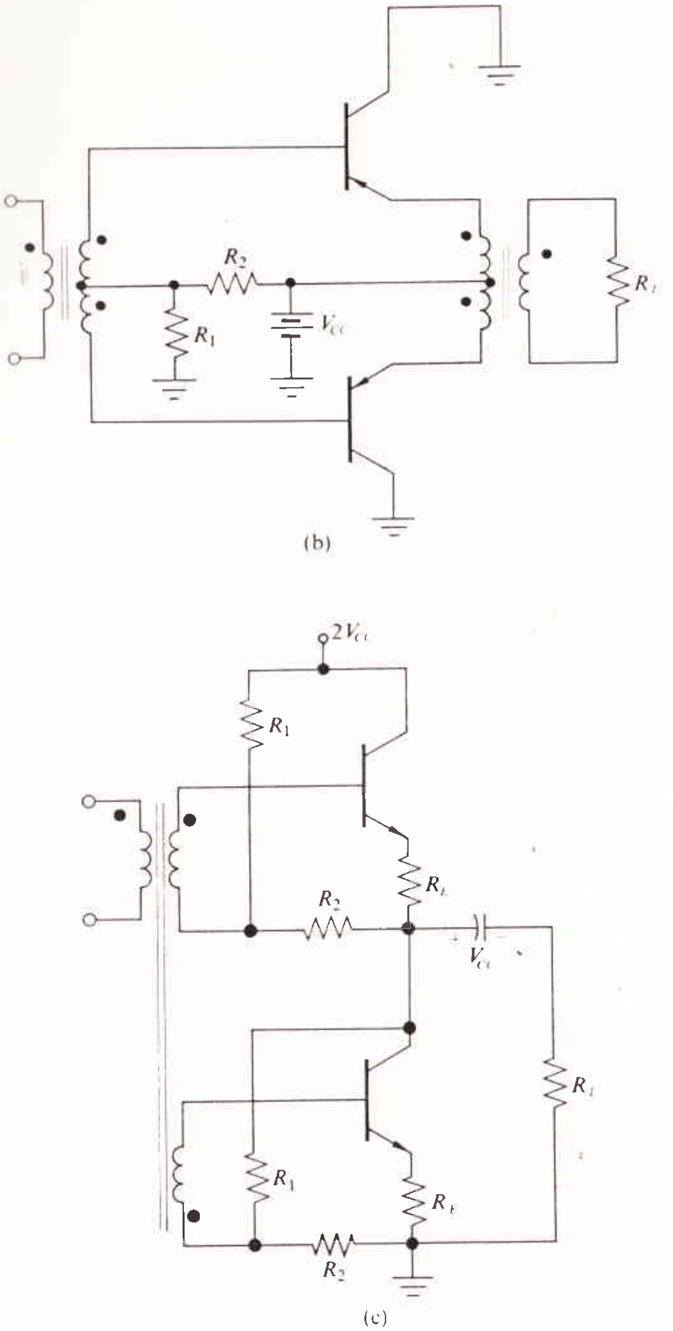


(a)

Figure 9.6–3

---

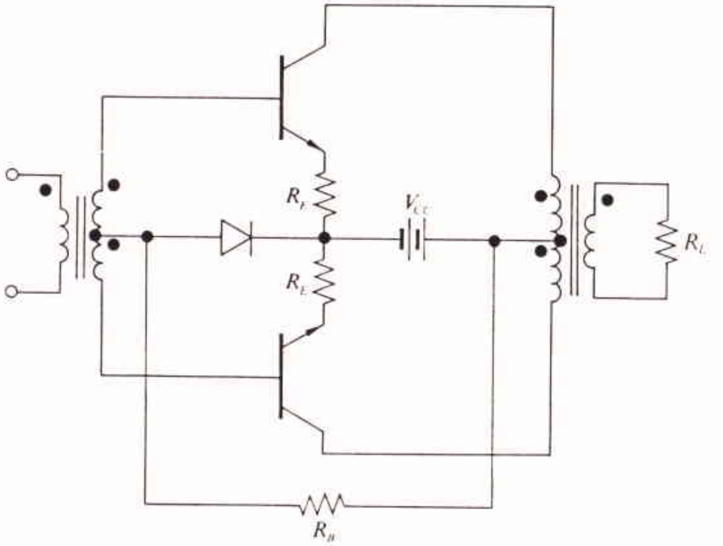
† The problem here is that high-power NPN transistors tend to be silicon, while high-power PNP units tend to be germanium. Even with different turn-on biases it is normally difficult to make such units track in output current over a wide range. The emitter-follower configuration eases these difficulties by minimizing the effect of the transistor.



(b)

(c)

Figure 9.6–3



(d)

Fig. 9.6–3 (a) Complementary symmetry Class B emitter follower. (b) Transformer-coupled Class B emitter follower. (c) Transformerless current-driven Class B stage. (d) "Classic" Class B current-driven stage. Except for the biasing arrangements and impedance levels, vacuum tubes could be substituted directly in this circuit.

The other three circuits of Fig. 9.6–3 suffer from poor bias stability unless some emitter resistance is included. [In Fig. 9.6–3(b) the winding resistance of the transformer may serve this purpose.] The problem is that, unless R_E is small with respect to the load (or the reflected load), it both consumes power and reduces the available output voltage swings. With a 5 A peak current, a 0.2 Ω resistor consumes 1 V of the available swing as well as 1.25 W of the output power.

All four circuits indicate a "turn-on" bias for the transistors. This bias is necessary to avoid a dead zone or cross-over zone where with small input signals both units would be off or so nearly off that the gain would be reduced. The trick here is to just minimize this distortion without causing excessive standby power dissipation.

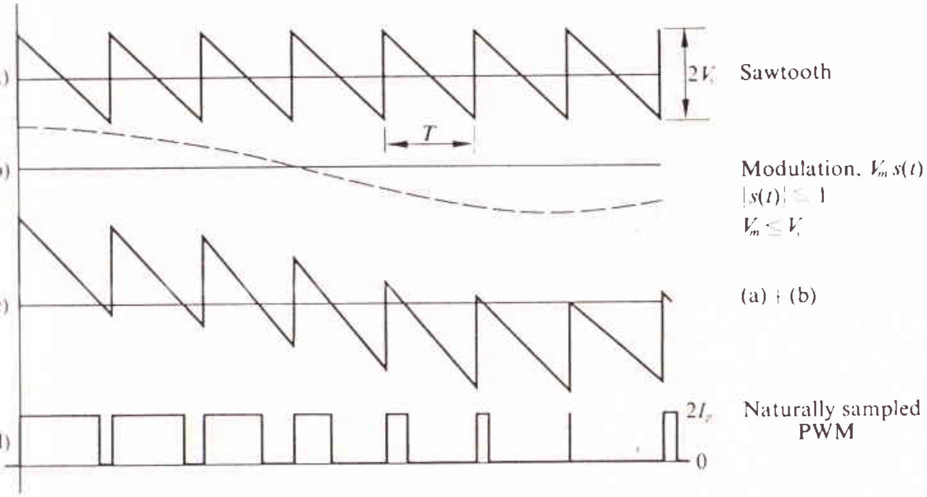
The circuit of Fig. 9.6–3(c) avoids an output transformer by using a large coupling capacitor. (The lower 3 dB point is at $1/RC$; hence if $R = 4\ \Omega$ and if $f_{\text{low}} = 20$ Hz is desired, then $C = 2000\ \mu\text{F}$.) The capacitor charges up to V_{CC} (we assume matched transistors), and the circuit is similar to the basic model of Fig. 9.6–1. Even though the top transistor has its emitter connected to the load, it is the collector current that flows through the load; hence β-matching is vital for this circuit. (The transistor operates as a current source, not as an emitter follower.)

The circuit of Fig. 9.6–3(d) shows a classic transformer-coupled Class B stage. Its low-frequency response is governed by the transformer-load combination, while its high-frequency limits may be set by transistor phase-shift mismatch or by transformer reactances.

BROADBAND CLASS D POWER AMPLIFIERS

ough most present-day power amplifiers are Class B double-ended units, it is mpting to examine the possibilities of extending the switching mode of operation the broadband case.

In the narrowband Class D case, the output circuit was tuned to the fundamental the switching frequency. For broadband operation we must switch at a frequency ıch higher than the top of our desired frequency band. The switching must be done such a fashion that the low-frequency average value is the desired output signal. ı output wideband bandpass filter will then remove the unwanted high-frequency rrier terms (and their sidebands) and the dc component.



g. 9.7-1 Generation of "naturally sampled" pulse-width modulation (trailing edge modula-n). Time origin at zero crossing of the sawtooth's downgoing ramp.

Fortunately, a switching signal of the proper type does exist. It is known as a aturally sampled" pulse-width-modulated signal. As developed in Fig. 4-12 260) and Eq. 220 (p. 284) of Rowe,† the signal illustrated in Fig. 9.7-1(d) has the llowing frequency components:

$$i(t)_{\text{PWM}} = I_p + I_p\frac{V_m}{V_s}s(t) - \frac{2I_p}{\pi}\sin\omega_0 t - \frac{2I_p}{\pi}\sin\omega_0[t + \tau s(t)], \qquad (9.7\text{-}1)$$

also has higher harmonics of ω_0 and sidebands around them where

$$\omega_0 = \frac{2\pi}{T}, \qquad |s(t)| < 1, \qquad \text{and} \qquad \tau = \frac{V_m T}{2V_s}.$$

ıus if $s(t)$ is band-limited and if ω_0 is sufficiently larger than the highest frequency in $s(t)$, then it will be possible to separate out an amplified signal proportional to $s(t)$ with a filter that rejects both dc and the "carrier" and its harmonics.

If $s(t) = \cos \mu t$, then the $\sin \omega_0[t + \tau s(t)]$ term will expand into a carrier at ω_0 and sidebands with amplitudes proportional to $J_n(\omega_0\tau)$ spaced at $n\mu$ around $\omega_0[J_n(\omega_0\tau)]$ is an ordinary Bessel function. Since $V_m \le V_s$ has been imposed as a restriction, it follows that $\omega_0\tau \le \pi$ and hence the values of $J_n(\pi)$ will give an indication of the maximum amplitude to be expected from the extreme sidebands†:

$$J_1(\pi) \approx 0.285, \quad J_2(\pi) \approx 0.485, \quad J_3(\pi) \approx 0.34, \quad J_4(\pi) \approx 0.016,$$
$$J_5(\pi) \approx 0.056, \quad J_6(\pi) \approx 0.016, \quad J_7(\pi) \approx 0.0038.$$

Thus if μ is at the top of the allowable "modulation" band and ω_0 is equal to or greater than 8μ, then the seventh sideband will interfere directly with the low-pass signal but the relative distortion will be only

$$\frac{2}{\pi}J_7(\omega_0\tau) = \frac{2}{\pi}J_7(\pi) = 0.0024,$$

or one-quarter of one percent.

Actually, if the circuit's low-pass filter does not fall off rapidly enough, the lower-order harmonics may be the limiting factor. For example, $J_3(\pi) \approx 100J_7(\pi)$; hence with $f_0 = 8f_m$ the low-pass filter must provide an attenuation of 1/100 or 40 dB only two decades above the maximum modulation frequency. To do this and simultaneously avoid attenuating the desired signal will require a very sophisticated low-pass filter. The problem is eased by raising the carrier frequency. The limits on such increases are device frequency limitations that prevent square-cornered operation and hence increase device dissipation and the difficulty of containing the RF interference caused by the harmonic-rich modulated carrier. (For example, one should avoid submultiples of the common IF frequency of 455 kHz.) A 120 kHz carrier should be satisfactory for ordinary non-high-fidelity amplification with $s(t)$ normally below 6 kHz.

A complete Class D audio amplifier is illustrated in Fig. 9.7-2. As in the narrowband case we may consider constructing a switching circuit to switch a constant current between two ends of our filter, or we may consider a switching circuit to effectively switch the input of the filter between two different voltages.

Figure 9.7-3 indicates the basic voltage-switching case. At first glance this circuit appears to be identical to the narrowband circuit of Fig. 9.5-1(a). This is a case where looks are deceiving; since this filter must function as a broadband bandpass filter, not as a narrowband filter, the switch is no longer driven in a square-wave fashion, and, as we shall see, the turn-on requirements for the switch in the broadband case are much more stringent than they are in the narrowband case.

When $R/L \gg 1/RC$, the poles in the input admittance of the filter are widely separated and its asymptotic magnitude-vs.-frequency plot is as shown in Fig. 9.7-4.

Harrison E. Rowe, *Signals and Noise in Communications Systems*. Van Nostrand, New York 965).

† For smaller values of $\omega_0\tau$ the higher-order Bessel functions decay much more rapidly than linearly; hence the maximum percentage distortion will occur at maximum modulation. See Chapter 11 for details.

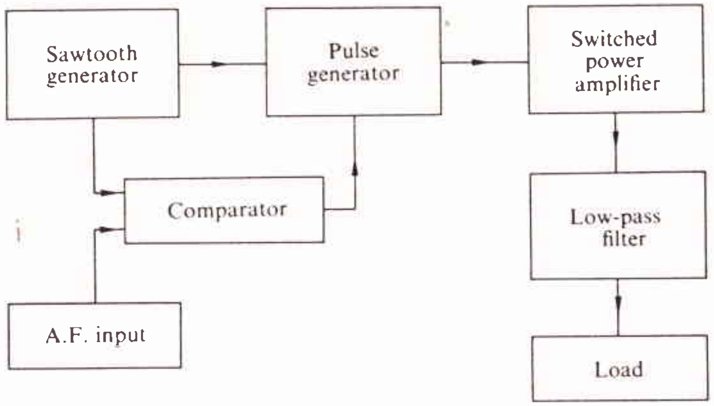


Fig. 9.7–2 Class D PWM broadband power amplifier.

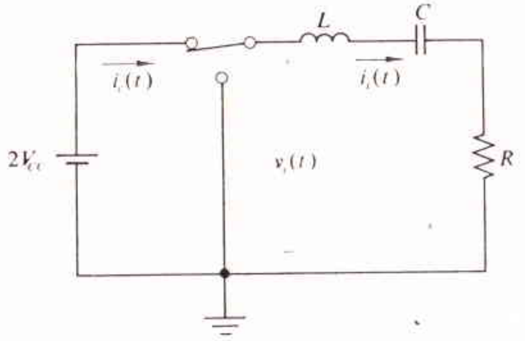


Fig. 9.7–3 Basic voltage-switching broadband Class D power amplifier.

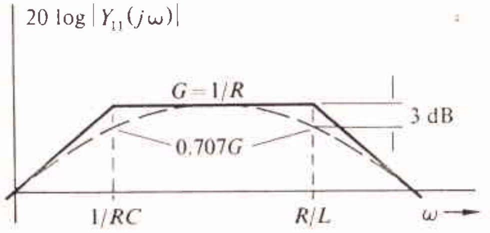


Fig. 9.7–4 Magnitude of the admittance of the filter of Fig. 9.7–3 ($R/L \gg 1/RC$).

If the switch is driven back and forth between the two terminals by a signal of the form shown in Eq. (9.7–1), then we may write

$$v_i(t) = V_{CC} + V_{CC}\frac{V_m}{V_s}s(t) - \frac{2V_{CC}}{\pi}\sin\omega_0 t - \frac{2V_{CC}}{\pi}\sin\omega_0[t + \tau s(t)]. \qquad (9.7\text{–}2)$$

If $s(t)$ is restricted to lie between $1/RC$ and R/L and if $\omega_0 \gg R/L$, then the capacitor will charge to V_{CC} and the network current will be

$$i_i = \frac{V_{CC}}{R}\frac{V_m}{V_s}s(t), \qquad (9.7\text{–}3)$$

where $V_m \leq V_s$ is an additional restriction. For example, if $V_m = V_s$ and $s(t) = \cos \mu t$, then

$$i_i = \frac{V_{CC}}{R}\cos \mu t.$$

As shown in Fig. 9.7–3, the switch has no dissipation; hence all the input power is converted into ac output power.

If the device (switch) has a peak current limitation $I_{\max}$, then

$$R \geq V_{CC}/I_{\max}.$$

If $R = V_{CC}/I_{\max}$, then $P_{ac} = V_{CC}I_{\max}/2$ and $P_{dc} = 2V_{CC}I_{dc}$; hence

$$I_{dc} = I_{\max}/4.$$

If the switch losses can be modeled completely as resistors R_s, then the efficiency becomes

$$\eta = \frac{R_L}{R_s + R_L}.$$

In this case the previous R is now $R_s + R_L$, where R_L is the actual load resistor.

Figure 9.7–5 indicates the filter input voltage and the current through V_{CC} for the case where $s(t)$ is a sine wave that lies between $1/RC$ and R/L. As the figure clearly indicates, the switch must carry current in both directions; hence for efficient operation without excessively large driving signals, the switches should be symmetrical devices (not vacuum tubes) capable of carrying current in the reverse direction.

If the signal is sinusoidal and driven so that the peak current through the switch is $I_{\max}$, then for the case where the switch is replaced by a battery, V_{sat},†

$$P_{ac} = \frac{(V_{CC} - V_{\text{sat}})I_{\max}}{2}, \qquad P_{dc} = 2V_{CC}I_{dc}, \qquad P_{\text{per switch}} = V_{\text{sat}}I_{dc}. \qquad (9.7\text{–}4)$$

Since $P_{dc} = P_{ac} + 2P_{\text{per switch}}$,

$$I_{dc} = \frac{I_{\max}}{4} \quad \text{and} \quad \eta = \frac{V_{CC} - V_{\text{sat}}}{V_{CC}}. \qquad (9.7\text{–}5)$$

---

† A battery is not really as good an approximation to the switch as a resistor is, since the switch current is not constant but varies over the cycle.

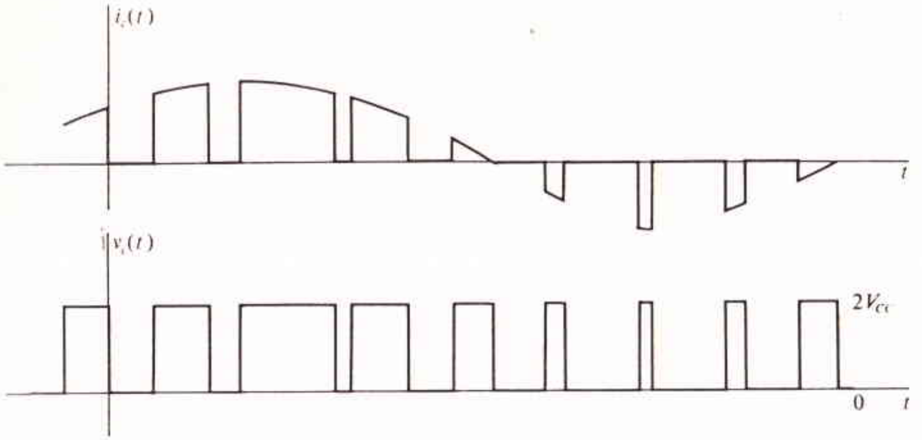


Fig. 9.7–5 Battery and upper switch current and filter input voltage when $s(t)$ is a sine wave ($V_m \approx V_s$).

This efficiency is, of course, exactly the value we achieved previously for a maximum-amplitude square-wave drive. The difference is that now we have a sinusoidal output. Even though our baseband may only extend to 5 kHz, we shall want a switching frequency of 100 kHz or more, and hence the problem of transition-time dissipation as discussed in Section 9.5 is still present.

Actually the situation is considerably worse in this case, since now the switch must be turned on and off while the current is at nearly its maximum value, while in the narrowband case theoretically the current was going through zero at every voltage-switching interval.

Figure 9.7–6 shows a possible mechanization of Fig. 9.7–3. When Q_1 is saturated, then Q_2 is on and Q_4 is off. With Q_2 saturated, Q_3 is turned on with (for the resistive values shown)

$$I_B \approx \frac{2(V_{CC} - V_{\text{sat}}) - V_{BE}}{66\,\Omega};$$

or with $V_{CC} = 10$ V and $V_{\text{sat}} = 0.5$ V for each of Q_1 and Q_2,

$$I_B \approx \frac{17.2}{66\,\Omega} = 260\text{ mA}.$$

If $\alpha_R \geq 0.8$, then the transistor Q_3 can be kept on for reverse currents of 1.3 A, since $I_B > (1 - \alpha_R)I_{CR}$ marks the edge of reverse saturation.

When Q_1 goes off, Q_2 and Q_3 go off and Q_4 is driven on with essentially the same base current as Q_3 had on the other portion of the cycle. With the values shown in Fig. 9.7–6 the load resistor should be $10/1.3 = 7.7\ \Omega$. For a smaller load the turn-on drive or the switching transistor reverse alphas would need to be increased.

Since the circuit always requires 260 mA of base current, it really would seem fair to charge it with $2V_{CC}I_B \approx 5$ W of additional power drain. It is true that the Class B stage also draws base current; however, in that case the base current is smaller and tracks the signal in amplitude instead of being constant.

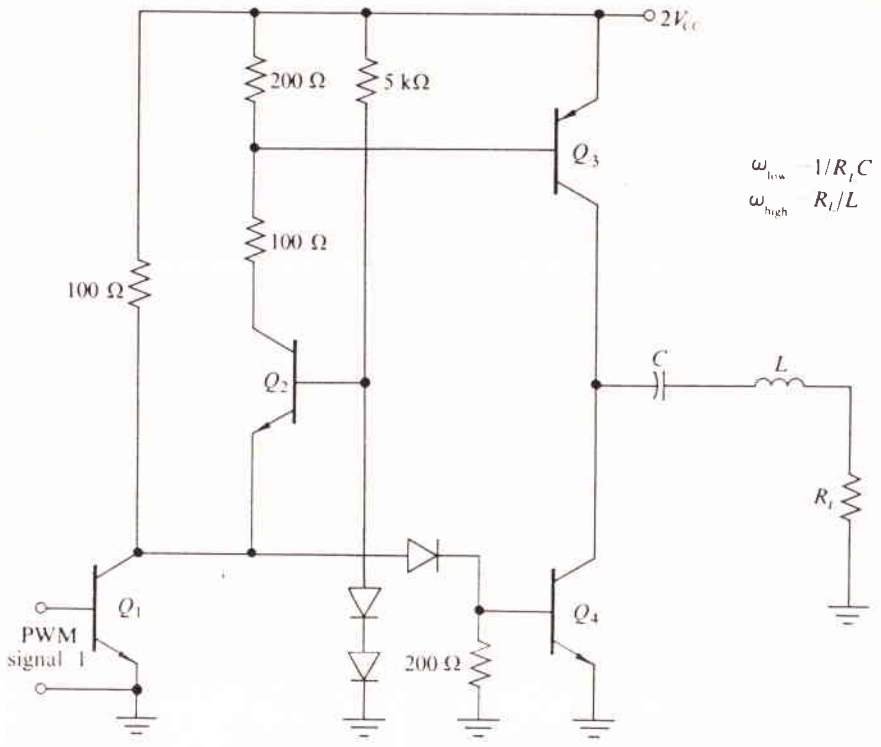


Fig. 9.7–6 Practical Class D pulse-width-modulated power amplifier.

Current-Switching Circuits

It would appear that one could construct a broadband current-switching circuit as the dual of the voltage case just discussed. The parallel RLC filter in this case would have $\omega_{\text{low}} = R/L$ and $\omega_{\text{high}} = 1/RC$. For a sinusoidal output from a configuration similar to Fig. 9.5–4 or 9.5–1(b), the individual "switch" currents would have the shape of v_i in Fig. 9.7–5 and a peak value of I_p (assuming a current source of $2I_p$). The "switch" voltages would have a sinusoidal "envelope," but would be zero whenever the switch was closed and I_p was flowing. This circuit would seem to have the advantage over the voltage switch of having to carry current in only one direction. It has the disadvantage of having to hold the transistor off for a reverse collector voltage equal to the peak output voltage.

The practical difficulty with such a circuit is that the true dual of the battery has not yet been invented. Thus one is forced to use a battery V_{CC} in series with an inductor L_{CC} as an approximation for a current source.

One now has difficulty properly relating L_{CC}, carrier frequencies (switching rates), and allowable modulation frequencies. Now, in addition to having the problem of separating the carrier frequency ripple from the modulation frequency in the output parallel *RLC* filter, one must separate these frequencies so that L_{CC} can be chosen to be an "open circuit" at the carrier frequency and a "short circuit" at the highest modulation frequency. This additional restriction tends to demand that the switching frequency be raised to be at least 100 times the highest modulation frequency, thus increasing the difficulty of obtaining a reasonable approximation to an ideal switching device.

Except in special cases in which one wants a constant-amplitude, high-efficiency, variable-frequency power amplifier, the current-switching broadband or Class D type of circuit would seem to be primarily of academic interest.

9.8 PRACTICAL POWER AMPLIFIERS

A number of practical problems arise when one attempts to construct a power amplifier. These problems include the provision of driving signals, the provision of proper bias, the design of circuits that protect themselves against accidental destruction, and the "matching" of the actual loads to the optimum impedance required by the device for efficient operation.

Let us consider first a fixed-amplitude Class C vacuum-tube power amplifier of the type considered in Section 9.4 and illustrated generally in Fig. 9.4–1. In this case we assume a bias and a drive and calculate the desired high-frequency resistive-load impedance. In the process of these calculations we also calculate the waveshape of the grid circuit current waveshape. From these waveshapes we can calculate the fundamental and dc values of current. Now with the assumed value of driving voltage we can calculate the effective input impedance of the circuit as well as a bias resistor that in conjunction with I_{G0} will supply the desired bias. If the bias is supplied completely by an I_{G0}-R_G combination, then when the drive circuit is turned off the bias is lost. For the case of Fig. 9.4–3 the tube dissipation at $V_{BB} = -12$ kV and $V_{CC} = 0$ is 66 kW, or more than six times the "normal" operating dissipation. Such a dissipation will undoubtedly lead to damage or destruction. One might choose to obtain enough of the bias from a fixed supply to prevent device destruction and the rest from a "grid leak" combination.

This self-biasing arrangement has the advantage that the bias adjusts itself to the driving signal in such a fashion as to tend to keep the output constant. That is, if the drive signal decreases, then so does the bias, or vice versa; hence the output current pulse shape changes much less than it would have if the bias were fixed and the same change in input level occurred.

Figure 9.8–1 illustrates two possible bias arrangements. In either case C_G should have an impedance at the operating frequency that is small enough so that I_{G1} flowing through it will cause a voltage drop of 1% or less of the desired driving voltage. Also, in either case, $V_{\text{partial}} + I_{G0}R_G$ should equal the desired bias voltage. Normally the circuit of Fig. 9.8–1(a) is to be preferred, since the other arrangement has an ac power loss in R_G that is not negligible and which must be supplied by the driver stage.

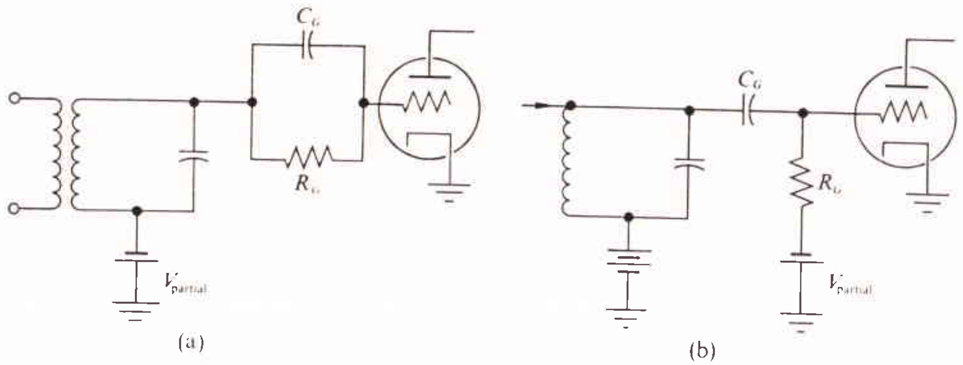


Fig. 9.8–1 Class C bias arrangements.

As an example, for the sinusoidal drive pulses illustrated in Fig. 9.4–4(a), if $V_{\text{partial}} = 1200$ V, then $R_G = 2200/0.45 = 4.9\ \text{k}\Omega$. The dc power loss in this resistor is $2200 \times 0.45 = 1000$ W. In the circuit arrangement of Fig. 9.8–1(b) there would be an additional ac loss of 3000 W in R_G. The power that must be supplied by the driver for Fig. 9.8–1(a) is $(5400 \times 0.85)/2 = 2300$ W, of which 540 W goes to "charging" V_{partial} and $2300 - 540 - 1000 = 760$ W is dissipated in heat at the grid circuit. Since the driver must supply 2.3 kW into a $5400/0.85 = 6350\ \Omega$ load, it must also be a power amplifier.

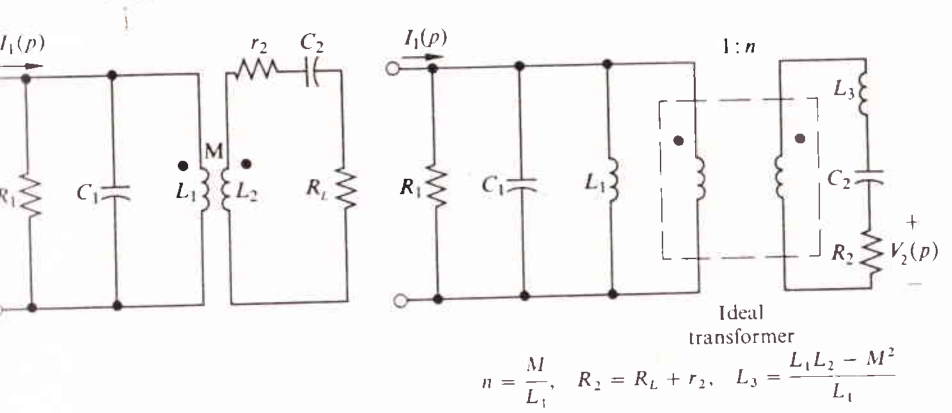
Drive Signals

Since i_G flows in pulses, it will contain a third-harmonic component; hence one could generate the required nonsinusoidal driving waveshape to supply the proposed current waveshapes of Fig. 9.4–4 by including a second tank circuit in series with those shown in the grid circuits of Fig. 9.8–1. If this second tank circuit is tuned to the third harmonic of ω_0 and is loaded so that I_{G3} causes a 540 V drop across it, then the desired grid waveshape will result (I_{G3} flows in the opposite direction to the input fundamental current from the driving stage). If one assumes that the current from the driver is purely fundamental by the time it reaches the secondary, then all the power supplied by the driving stage is at the fundamental frequency.

A "Class C" transistor stage is usually run with zero bias or with at most a small emitter resistor; hence the input circuit calculations are much simpler than those of the vacuum tube case. In general, with transistor circuits the fundamental input impedance of a power amplifier will be very small, perhaps only a few ohms; thus an impedance step-down circuit or transformer will be necessary between the driver and the power amplifier.

Coupling Networks

It is relatively simple to say what the coupling network should do, but somewhat more cumbersome to explain how to make it do what is required. Ideally, the coupling twork should "transform" the actual load to the desired resistive value at the perating frequency while presenting zero input impedance at all harmonic fre-uencies. At the same time the network should have a transfer function that passes e desired modulation passband without causing amplitude or phase distortion while ffering infinite attenuation to all harmonic frequencies. To some extent these quirements may be contradictory. Certainly they must all be checked individually, nce satisfying one of them does not guarantee the satisfaction of the others.



$$n = \frac{M}{L_1}, \quad R_2 = R_L + r_2, \quad L_3 = \frac{L_1 L_2 - M^2}{L_1}$$

(a) (b)

Fig. 9.8–2 Output coupling circuit. (a) Coupling circuit. (b) Possible circuit model.

The classic coupling network for vacuum-tube power amplifiers provides an xample of an almost ideal coupling situation. The circuit and a useful model for it re shown in Fig. 9.8–2. This particular model is chosen from the multitude of ossible models because it leads to a simple presentation of both Z_{in} and the transfer mpedance.

If the admittance to the left of the transformer in Fig. 9.8–2(b) is called $Y_1(p)$ while the admittance to the right is called $Y_2(p)$, then it is apparent that

$$Z_{\text{in}}(p) = \frac{1}{Y_1(p) + n^2 Y_2(p)} \quad \text{and} \quad \frac{V_2(p)}{I_1(p)} = \frac{n Y_2(p) R_2}{Y_1(p) + n^2 Y_2(p)}. \tag{9.8–1}$$

Both the input and the transfer impedance have identical poles. The transfer im-pedance has two zeros at the origin and a scale factor of $\omega_0^2 M R_2/L_3$, while the input impedance has complex zeros at the roots of $p^2 + p(R_2/L_3) + 1/L_3C_2$ and a single zero at the origin. The scale factor for this case is $1/C_1$.

The poles for either function may be found by the root locus technique. From Eq. (9.8–1) the poles of either function are at the roots of

$$0 = 1 + \frac{p^2 n^2}{L_3 C_1\left(p^2 + \dfrac{p}{R_1 C_1} + \dfrac{1}{L_1 C_1}\right)\left(p^2 + p\dfrac{R_2}{L_3} + \dfrac{1}{L_3 C_2}\right)}. \tag{9.8–2}$$

In order to concentrate on the important aspects of the circuit, we make the following very reasonable assumptions or approximations:

a) Since the unloaded Q of the primary alone should be equal to 100 or more, we assume, for the purpose of locating the first pole pair in the denominator of Eq. (9.8–2), that R_1 approaches infinity. (To calculate the network loss and the network transfer efficiency, we must reinstate the actual R_1.)
b) With respect to the upper complex pole pair we neglect the angles from the lower poles and zeros. If $R_2/2L_3 \geq \omega_0/4$, that is, if the loaded Q of the mythical secondary exceeds 2, and if we are interested in the loci from and after the poles coalesce, then the maximum angular error is 14° or 7° from each upper pole. Certainly in an initial consideration this neglect is justified in terms of obtaining a simpler explanation.

With these assumptions, if the upper complex poles are initially lined up horizontally, that is, if

$$\frac{1}{L_1 C_1} + \left(\frac{R_2}{2L_3}\right)^2 = \frac{1}{L_3 C_2} \tag{9.8–3}$$

or

$$\frac{1}{L_3 C_2} = \omega_0^2\left[1 + \left(\frac{R_2}{2\omega_0 L_3}\right)^2\right], \tag{9.8–4}$$

where $\omega_0^2 = 1/L_1C_1$, then as n increases they will move toward each other and coalesce at $R_2/4L_3$ from the j-axis; they will then move apart along a line parallel to the j-axis. Figure 9.8–3 illustrates the simplified locus.

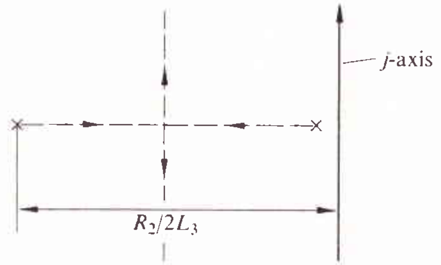


Fig. 9.8–3 Simplified locus of upper pole pair.

UNIVERSIDAD DE LA REPUBLI
FACULTAD DE INGENIE
DEPARTAMENTO DE DOCUMENT
Y BIBLIOTECA
MONTEVIDEO - URUGU

If we restrict ourselves to values of n that cause coalescence or vertical splitting, then from the root locus magnitude condition and from Eq. (9.8–2)

$$\frac{n}{2\sqrt{L_3 C_1}\, d} = 1, \tag{9.8–5}$$

where d is the distance from the original right-hand pole (actually from either original pole) to the desired coupled pole positions.

Two positions would appear to be of primary interest. In one, the poles have just coalesced. In the second significant position, the poles have moved apart vertically by 45° so as to yield a maximally flat transfer function. The results for these two cases are summarized in Table 9.8–1, where R_{in} is the desired input impedance at ω_0.

Table 9.8–1

	Coalescence	Maximally flat
d	$R_2/4L_3$	$R_2/2\sqrt{2}L_3$
$R_{\text{in}}R_2$	$4L_3/C_1$	$2L_3/C_1$
M/L_1	$\sqrt{R_2/R_{\text{in}}}$	
L_1	$1/\omega_0^2 C_1$	
L_2	$L_3 + M^2/L_1$	
C_2	$\dfrac{1}{\omega_0^2 L_3\left[1 + \left(\dfrac{R_2}{2\omega_0 L_3}\right)^2\right]}$	
η_1	$\dfrac{R_1}{R_1 + R_{\text{in}}}$	
η_2	$\dfrac{R_2 - r_2}{r_2}$	
k	$M/\sqrt{L_1 L_2}$	
$\left.\dfrac{V_2}{I_1}\right\|_{\omega=\omega_0} = Z_{21}(j\omega_0)$	$\sqrt{R_{\text{in}}R_2}$	

Table 9.8–2 indicates the relative magnitudes of Z_{in} and Z_{21} as a function of $\delta\omega$, where $\delta\omega$ is measured from $\omega_0 = 1/\sqrt{L_1C_1}$. Note that by the second harmonic we are assuming that the responses for the two cases are identical. This is not an exact result, but is a very good initial approximation. It is apparent that the maximally flat case is to be preferred whenever it is possible, since it gives a better "in-band" characteristic without causing any appreciable difference in its response to harmonics.

To place these results in perspective, let us consider a numerical example. Assume that

$$R_{\text{in}} = 2000\,\Omega, \qquad R_2 = 50\,\Omega, \qquad \omega_0 = 10^7, \qquad \text{and} \qquad Q_1 = Q_2 = 100.$$

Now to keep the primary circuit losses down to 5% of the power passing through it,

$$\frac{R_1}{R_1 + R_{\text{in}}} = 0.95 \qquad \text{or} \qquad R_1 = 20R_{\text{in}} = 40\ \text{k}\Omega.\dagger$$

---

† Actually R_{in} should include R_1. So long as $R_1 \gg R_{\text{in}}$, the separation shown is reasonable.

Table 9.8–2

$4\delta\omega L_3/R_2$	Coalescence		Maximally flat	
	$\|Z_{\text{in}}\|/R_{\text{in}}$	$\|Z_{21}\|/Z_{21}(0)$	$\|Z_{\text{in}}\|/R_{\text{in}}$	$\|Z_{21}\|/Z_{21}(0)$
0.0	1.00	1.00	1.00	1.00
±0.5	0.83	0.80	1.02	0.99
±1.0	0.56	0.50	1.00	0.89
±1.5	0.38	0.31	0.83	0.67
±2.0	0.28	0.20	0.63	0.44
+nth harmonic	$\dfrac{n}{n^2-1}\,\dfrac{1}{\omega_0 C_1 R_{\text{in}}}$	$\left(\dfrac{n}{n^2-1}\right)^2 \times \left(\dfrac{2}{\omega_0 C_1 R_{\text{in}}}\right)^2$	$\dfrac{n}{n^2-1}\,\dfrac{1}{\omega_0 C_1 R_{\text{in}}}$	$\left(\dfrac{n}{n^2-1}\right)^2 \times \left(\dfrac{2}{\omega_0 C_1 R_{\text{in}}}\right)^2$

If $Q_1 = 100$ and $R_1 \geq 40\,\Omega$,

$$L_1 \geq 40\ \mu\text{H} \qquad \text{for } \omega_0 = 10^7.$$

If we choose $L_1 = 40\ \mu$H, then

$$C_1 = 1/\omega_0^2 L_1 = 250\ \text{pF}$$

and

$$M = L_1\sqrt{R_2/R_1} = \sqrt{40}\ \mu\text{H} = 6.35\ \mu\text{H}.$$

From Table 9.8–1 (the second line) we see that the difference between the two pole positions rests in L_3 (and also, therefore, in L_2 and C_2). The results for the two cases are shown in Table 9.8–3.

The maximally flat case has a slightly lower circuit efficiency ($\eta_{\text{circuit}} = \eta_1\eta_2$), however, it also has a lower coefficient of coupling, k, and a smaller value for C_2. A practical problem remains as to whether the desired coil combination is physically possible. In this case, if L_2 is placed over the end of L_1, then the required amount of coupling should be possible.

Table 9.8–3

	Coalescence	Maximally flat
L_3	6.25 μH	12.5 μH
L_2	7.25 μH	13.5 μH
C_2	1380 pF	770 pF
k	0.374	0.275
η_2	0.985	0.973
η_{circuit}	0.935	0.925

We may now examine the relative input and transfer impedance at the second armonic to see how effective the network is in suppressing harmonics:

$$\frac{|Z_{\text{in}}(2j\omega)|}{R_{\text{in}}} = \frac{2}{3}\frac{1}{5} = \frac{2}{15}, \qquad \frac{|Z_{21}(2j\omega)|}{|Z_{21}(j\omega)|} = \frac{4}{9}\frac{4}{25} = \frac{16}{225}.$$

hese results indicate that in a case similar to the one shown in Fig. 9.4–4(a), where $_2/I_1 = 0.72$, the second-harmonic voltage across the input of the tank circuit will e approximately 10% of the fundamental voltage, while the second-harmonic utput voltage will be 5% of the fundamental voltage.

Since the input harmonic voltage is shifted in phase by nearly 90° with respect o the fundamental, one can approximate the input tank voltage (this voltage adds o V_{CC} to make the device voltage) by

$$V_1 \cos \omega_0 t + 0.1 V_1 \sin 2\omega_0 t.$$

he effect of such a tank voltage on the operating path, with a pure sinusoidal driving aveshape, is shown in Fig. 9.8–4.

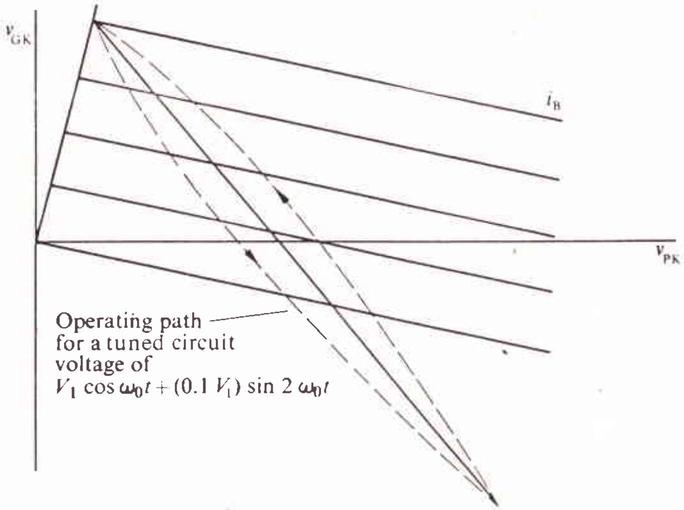


Fig. 9.8–4 Effect of 90° shifted second harmonic on the operating path.

Several things are apparent from this new path. One is that the output current ulse shape is different from that found previously and is nonsymmetrical. Because he pulse shape is different, the previous calculations will be somewhat in error. ecause the current pulse is not symmetrical, there will be a phase difference between he assumed tank voltage and the resultant fundamental from the current pulse. ormally one does not want to pursue this calculation any further around the loop. hat one does want to do is restrict the principal input harmonic voltage to the order f 10% of the fundamental.

If one wants to reduce the harmonics, one should increase C_1; however, for a fixed ω_0, increasing C_1 will reduce L_1 and cause the primary circuit efficiency to suffer. For the case shown, we have presented a reasonable compromise between the two positions.

The key term in harmonic reduction is $\omega_0 C_1 R_{\text{in}}$, which is the "loaded" Q of the primary circuit. We have used a value of 5. A value of 10 would reduce the output harmonic voltage by a factor of 4, but would do so at the expense of network efficiency. For low-power circuits one often does sacrifice the output power for the sake of improved filtering.

If a push-pull output stage is employed, the second-harmonic terms should ideally cancel. Under these circumstances one might halve C_1 without making the harmonic spreading of the operation path much worse. This would allow one to double L_1 and, if Q_1 can be maintained, to halve the primary losses. (In a megawatt transmitter 5% is 50 kW; hence reducing the primary losses from 5% to 2.5% is a worthwhile endeavor.)

In closing the discussion of this network, it should be pointed out that the values for the midband Z_{in} and transfer impedance can be obtained almost by inspection if both sides of the transformer are assumed to be tuned to resonance. This approach, combined with the idea that the harmonic content of both the input voltage and the output voltage depends essentially on the loaded Q of the primary circuit, i.e., on $\omega_0 C_1 R_{\text{in}}$, has been in use for perhaps fifty years. What has not been widely appreciated is that a technique like the root locus method lets one shape the "in-band" response in a relatively simple fashion.

Pi Network

Another network that is widely used in the output of power amplifier stages is the pi network of Fig. 9.8–5. In this figure R_{C1}, R_{C2}, and r are all loss elements that must be considered in calculating network efficiency but that may, one hopes, be neglected in estimating the network frequency response.

Unfortunately, the pi network is not as amenable to a general treatment as is the coupled coil circuit just considered. The basic conceptual problem is that the input impedance of the pi network has a pair of complex zeros and a pair of complex poles rather than the two sets of complex poles found in the coupled coil case. As we shall

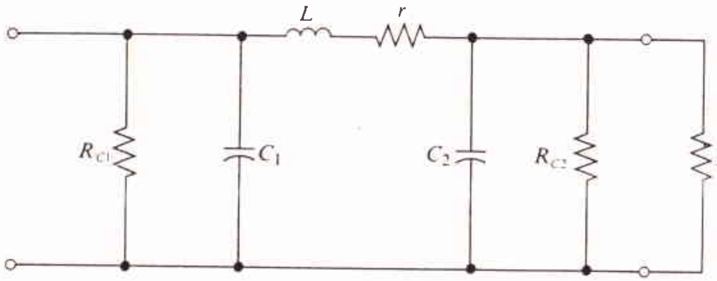


Fig. 9.8–5 Pi coupling network.

e, the upper pole and zero are usually close enough together and low enough in eir values of Q that the simple-minded approximations normally used for dealing ith such circuits tend to break down. In addition the presence of both a complex pole nd a complex zero near the operating frequency makes the passband nonsymmetrical nd forces one to reexamine such concepts as "bandwidth."

We shall see that the transfer impedance has the same poles as the input im-edance, but is lacking the complex zeros. However, since the center of the passband rarely situated directly "under the pole," the transfer impedance, $Z_{21}(p)$, is also onsymmetrical about the center frequency.

Before we consider the general case, let us examine an approach to calculating e input impedance of the circuit of Fig. 9.8–5 in which we assume that all the com-onents have known values. The method consists of combining impedances in an ternating parallel and series fashion as one starts with the load and works back ward the input of the network. That is, one combines R_{C2} and R_L to form R_2 and en, at the operating frequency, converts the parallel C_2-R_2 combinations into an uivalent series R_sC_s form:

$$C_s = \frac{Q_2^2 + 1}{Q_2^2} C_2, \qquad R_s = \frac{R_2}{1 + Q_2^2}, \qquad Q_2 = \omega_0 C_2 R_2. \tag{9.8–6}$$

or example, if $\omega_0 C_2 R_2 = 3$, then

$$C_s = (10/9)C_2 \qquad \text{and} \qquad R_s = R_2/10.$$

ne then combines R_s and r and $j\omega L$ and $1/j\omega C_s$. If the input impedance is to be sistive, then $\omega L > 1/\omega C_s$ is necessary so that the result of the combination is an uivalent series L-R combination. This combination is then converted back into arallel form and the inductive part is resonated out, with C_1 leaving the parallel ombination of R_{C1} and R_p as R_{in}:

$$R_p = (1 + Q_s^2) r_s, \qquad L_p = L_s \frac{Q_s^2 + 1}{Q_s^2}, \qquad Q_s = \frac{\omega_0 L_s}{r_s}, \tag{9.8–7}$$

here

$$r_s = R_s + r, \qquad L_s = L - \frac{1}{\omega_0^2 C_s}.$$

From this approach we see that the network power transfer efficiency will be

$$\eta_{\text{circuit}} = \frac{R_{C1}}{R_{C1} + R_p} \frac{R_s}{R_s + r} \frac{R_{C2}}{R_{C2} + R_L}. \tag{9.8–8}$$

ence, for high efficiency, $R_{C2} \geq 100R_L$, $R_s \geq 100r$, and $R_{C1} \geq 100R_p$ are all esirable relationships.

Now that the loss terms have been considered, let us neglect them in order to mplify the algebra while we examine the network frequency response. Consider two numerical examples (in both cases $C_1 = C_2$):

$$\begin{array}{rlcrl}
\omega_0 C_2 R_2 &= 3, & \qquad & \omega_0 C_2 R_2 &= 3, \\
Q_2 &= 3, & & Q_2 &= 3, \\
R_s &= R_2/10 & & R_s &= R_2/10, \\
\omega_0 L/R_2 &= 6/10 & & \omega_0 L/R_2 &= 1/3, \\
Q_s &= 3, & & Q_s &= 1/3, \\
R_p &= R_2, & & R_p &= R_2/9, \\
\omega_0 C_1 R_{\text{in}} &= 3, & & \omega_0 C_1 R_{\text{in}} &= 1/3.
\end{array}$$

In one case we have no impedance change, while in the other we have a 9/1 impedance step-down. If we started with a 50 Ω load, then the step-down case would be useful for a transistor power amplifier.

As we shall see shortly, by choosing a larger value for $\omega_0 L/R_2$ in the first case, and by changing C_1 in an appropriate fashion, we could have obtained a large impedance step-up ratio rather than the 1/1 case obtained.

In order to see how such a wide range of input impedances is possible, as well as to begin to understand the frequency response of the network, let us write the expressions for Z_{in} and Z_{21}, assuming lossless elements in both cases:

$$Z_{\text{in}}(p) = \frac{p^2 + \dfrac{p}{R_2 C_2} + \dfrac{1}{LC_2}}{C_1\left(p^3 + \dfrac{p^2}{R_2 C_2} + p\dfrac{C_1 + C_2}{LC_1C_2} + R_2C_2LC_1\right)}. \tag{9.8–9}$$

$$Z_{21}(p) = \frac{\dfrac{1}{LC_1C_2}}{\text{same poles as } Z_{\text{in}}(p)}. \tag{9.8–10}$$

If we define $\omega_2^2 = 1/(LC_2)$, $C_2 = NC_1$, and $Q_2 = \omega_0 C_2 R_2$ then $Z_{\text{in}}(p)$ may be rewritten as

$$Z_{\text{in}}(p) = \frac{p^2 + p\omega_0/Q_2 + \omega_2^2}{C_1(p^3 + p^2\omega_0/Q_2 + p[1 + N]\omega_2^2 + \omega_2^2 N\omega_0/Q_2)}. \tag{9.8–11}$$

where ω_0 is the desired operating frequency—that is, the frequency at which $Z_{\text{in}}(j\omega)$ is real and has the desired magnitude.

If one knows the circuit elements, then the calculation of $Z_{\text{in}}(j\omega)$ is straightforward, although tedious if done by hand rather than by a stored program in a minicomputer. The problem is that—rather than knowing the circuit elements—one knows the frequency and the impedance—transformation ratio, and wishes to choose the elements so as to perform the transformation in some optimum manner.

To get some further feeling for the problem, we might let Q_2 approach infinity (that is, let the effective load resistance R_2 approach infinity) in Eq. (9.8–11). The esultant equation would have complex zeros at $\pm j\omega_2$ and complex poles at $j\sqrt{1+N}\omega_2$, as well as a pole at the origin. Now from root-locus consideration, t may be shown that as R_2 decreases from infinity the actual poles and zeros will move nto the left-hand plane, initially along paths that are perpendicular to the j-axis. The Q of the complex zeros is just Q_2, while—again from root-locus considerations—the Q of the complex poles may be shown to be on the order of 1.5 to 2.5 times Q_2.

If one is trying to obtain an impedance step-up, then he would expect to operate in the vicinity of the upper complex pole, while if he wishes an impedance step-down, he would expect to operate more nearly "under the zero" of Z_{in}.

As we shall see, for the values of Q_2 in the range from 1 through 5, neither the step-up nor the step-down cases will be particularly symmetrical with respect to frequency. In the impedance step-down case, one has an additional complication in that—unless the value of $N = C_2/C_1$ is chosen carefully—one will find that the complex pole in the impedance functions will end up near one of the harmonics of the desired operating frequency. Since, in power-amplifier usage, the network input current is normally very rich in harmonic components, the result will be a large voltage component at the undesired harmonic frequency. For example, if $N = 3$ ($C_2 = 3C_1$), then the complex pole occurs near $2\omega_0$ and both the input impedance and the transfer impedance may have second harmonic terms that greatly exceed the fundamental magnitudes. As we saw in connection with Fig. 9.8–4, such a situation leads to undesirable operating conditions for the output device. It also leads to undesired interference terms in the network's output voltage.

This problem of the large harmonic component is reduced by getting the complex pole away from the harmonic frequency. This end is accomplished by reducing the value of N (that is, by increasing C_1 with respect to C_2). For any given Q_2 there is a limit to the amount that C_1 can be increased and still allow one to obtain a real input impedance at the desired operating frequency. If $Q_1 = \omega_0 C_1 R_2$, then this limit may be expressed as

$$Q_1 \le \frac{Q_2+1}{2}, \qquad N \ge \frac{2Q_2}{Q_2^2+1}.$$

While for most combinations of Q_2 and N there are two operating frequencies at which Z_{in} is resistive—the step-up and the step-down cases—in the limiting case these frequencies coalesce into a single point.

Table 9.8–4 relates several of the circuit parameters, at this limiting value of N, to various values of Q_2.

The value of the input impedance at the second harmonic is found by evaluating the input impedance with the values of components and frequencies specified by the earlier values in the table. While the exact impedance expression is somewhat cumbersome, for all values of $Q_2 \ge 1$ it may be closely approximated by

$$Z_{\text{in}}(2j\omega_0) \approx -j\left(\frac{R_2}{Q_1}\right)\left[\frac{3Q_2+4}{6Q_2+4}\right]. \tag{9.8-12}$$

Table 9.8–4 Pi network parameter when $N = C_2/C_1$ has its minimum allowable value

$Q_2 = \omega_0 C_2 R_2$	$\omega_0 L/R_2$	$Q_1 = \omega_0 C_1 R_2$	R_{in}/R_2	ω_0/ω_2	$\lvert Z_{\text{in}}(2j\omega_0)\rvert/\lvert Z_{\text{in}}(j\omega_0)\rvert$
1	1.0	1.0	1/1	1.0	0.724
2	0.60	2.5	1/2.5	1.10	0.635
3	0.40	5.0	1/5	1.10	0.590
4	0.29	8.5	1/8.5	1.085	0.575
5	0.23	13.0	1/13	1.075	0.560
6	0.19	18.5	1/18.5	1.062	0.550

For $Q_2 = 1$ the difference between the exact and the approximate expressions is less than 5 degrees in angle and less than 3.5% in magnitude.

From Eq. (9.8-12) and Table 9.8–4, we see that even if Q_2 approaches infinity, the magnitude of the second harmonic term will still be one-half of the magnitude of the term at the operating frequency. From our previous calculations we are aware that this may well lead to a device output voltage component at the second harmonic equal to 20–30% of the fundamental component. Even if the device output voltage has no controlling effect on the output current pulse (an ideal transistor, for example), this large harmonic component will shift the device operating path into a higher dissipation region, as indicated by Fig. 9.8–4. Hence it is not desirable.

A rational way to remove the second harmonic term from both the input and the transfer impedances is to parallel C_1 by a series-tuned, high-Q, low-loss LrC circuit that is resonant at $2\omega_0$. To be effective, this circuit must have a resonant impedance at $2\omega_0$ that is small with respect to the desired output impedance at ω_0. In many high-powered transistor circuits, this desired impedance is only a few ohms; hence the series trap must have an impedance less than a few tenths of an ohm.

Inclusion of such a series circuit will move the poles of Z_{in} as well as introducing a new set of poles above $2\omega_0$. If the Q of the trap is high enough ($Q > 100$ is desirable), then these new poles will be close enough to the $2\omega_0$ so that they will not cause problems at the third harmonic, while their effect at the operating frequency will be only second order.

The problem remains of obtaining enough of a handle on the "bandwidth" of the circuit to be able to specify it in an intelligent manner.

To see the effect of variations in frequency, we plotted the magnitude and phase of $Z_{\text{in}}(j\omega)$ vs. ω/ω_0 for the case in which $Q_2 = 3$, $N = 0.60$, and hence $(\omega_0 L/R_2) = 0.4$. That is for the limiting 5/1 step-down case of Table 9.8–4. With these values, a 50-ohm load is transformed into a 10-Ω load when $\omega = \omega_0$, into $7.45\,\Omega\angle -6°$ when $\omega = 0.95\omega_0$, and into $13.8\,\Omega\angle -2.5°$ when $\omega = 1.05\omega_0$. Thus a 5% shift in frequency causes roughly a 30% shift in the magnitude of the input impedance, and only a relatively small shift in the phase angle.

Higher values of Q_2 will move the complex zeros closer to the j-axis, and hence will allow larger step-down ratios, while keeping the impedance at $2\omega_0$ from becoming excessive. For a given value of Q_2, reducing $\omega_0 L/R_2$ from the values in Table 9.8–4 will increase the step-down ratio at the expense of an increase in the relative second-harmonic component. If a very large step-down is required, then a conflict may arise between bandwidth and step-down ratio. Two cascaded pi sections may provide a solution to this problem. However, the extra reactances in the circuit will also provide extra poles that must be steered clear of the harmonics of the operating frequency.

Smith Chart

A practical device that yields some insight into this matching problem is a modified version of the transmission line chart known as the Smith chart. If one normalizes all reactances (or conductances) with respect to the load, which in our case is R_2, then the chart, illustrated in Fig. 9.8–6, allows one to see by inspection what component ratios are allowable. It also allows the simple inclusion of losses from the network's inductances and capacitances, as well as allowing one to calculate the effects of variations in frequency on the impedance.

The mechanism of the chart that is useful is that a "reflection" of a normalized point through the center of the chart performs the parallel-series and series-parallel conversions of Eqs. (9.8–6) and (9.8–7).

The usual version of the Smith chart has resistive (or conductive) circles centered on a vertical bisector of the chart and reactive circles centered off to the sides of the chart. To be useful for our purposes the chart needs several "reflection" circles added, as shown in Fig. 9.8–6. The circle labeled 1 is a reflection into the upper half of the chart of the $(R/Z_0) = 1$ circle from the lower half of the chart. Since we have normalized $\omega_0 C_2$ with respect to G_2 the first point on the chart always lies in the lower right-hand quadrant at the intersection of the 1.0 "resistive" circle and the "reactive" circle equal in numerical value to $Q_2 = (\omega_0 C_2/G_2)$. The left-hand half of the reflected $(R/Z_0) = 1$ circle always contains the second point as the "reflection" of the first point through the center of the chart. Now for the step-down case we plot in the upper left quadrant the reflection half-circle(s) corresponding to the desired step-down ratio(s). That is, for a 5/1 impedance step-down the 5 circle tangent to the bottom of the chart is reflected to the top of the chart.

Now unless we can travel around (to the right) on the resistive circle from the first reflection point in the upper left-hand quadrant and intersect the reflected step-down circle, then this combination of Q_2 and step-down are not allowable. If the step-down value cannot be reduced, then either Q_2 must be increased or the step-down must be obtained as a two-stage process.

A numerical example may help to clarify the use of the modified chart. First we normalize the components of Fig. 9.8–5 to R_2 or G_2.

$$\omega C_2' = \omega C_2 R_2, \qquad \omega L' = \omega L/R_2, \qquad \omega C_1 R_2 = \omega C_1',$$

$$G_{\text{in}}' = G_{\text{in}}/G_2, \qquad \text{and} \qquad G_2' = G_2/G_2 = 1.$$

As a numerical example related to our previous calculations, consider the case in which $Q_2 = \omega_0 R_2 C_2 = \omega C_2' = 3$ and for the case of lossless components find the values (normalized) of L and C_1 that will provide a 4/1 impedance step-down.

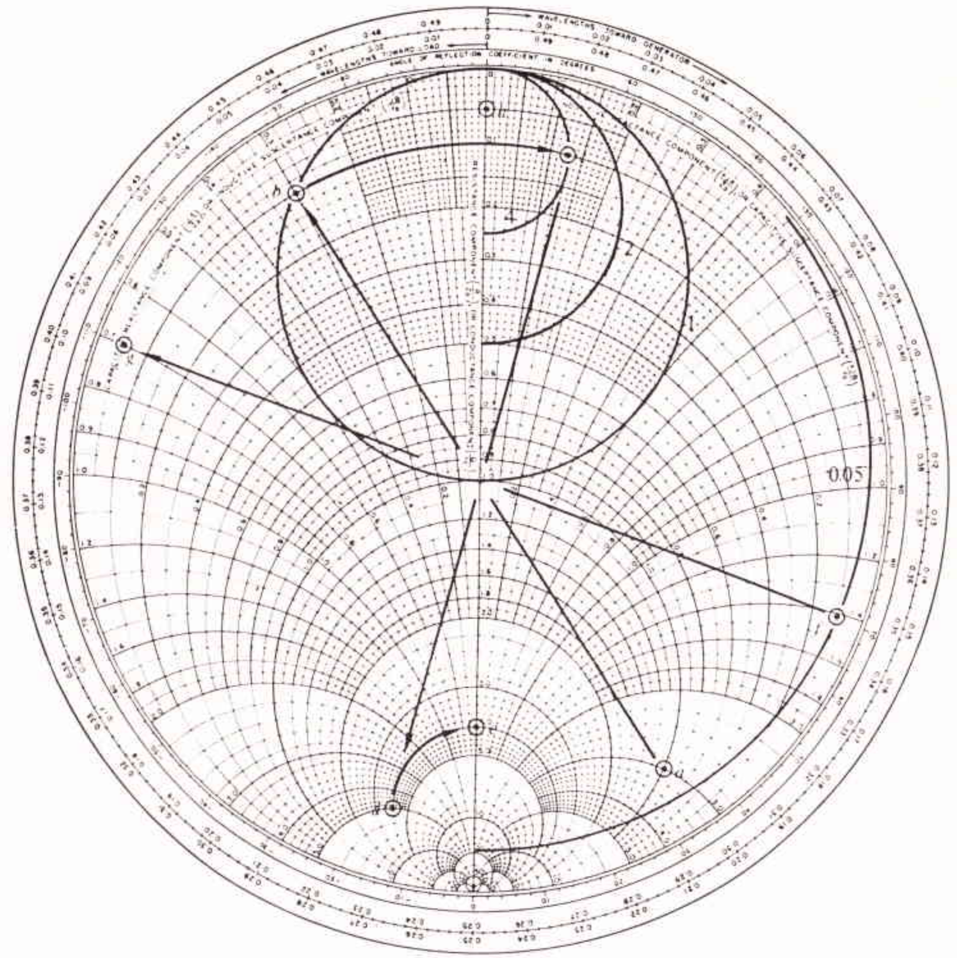


Fig. 9.8–6 Modified Smith chart. Added reflection circles aid in the easy design of pi matching networks.

The normalized $\omega_0 C_2 - G_2$ conductance is $1 + j3$, which is shown as point a in Fig. 9.8–6. The parallel-to-series conversion is performed by reflection through the center of the chart to point b, which has the normalized series impedance of $0.1 - j0.3$. Now we proceed to the right along the 0.1 resistive circle until we strike the 4/1 step-down circle at point c. We then reflect through the center to point d ($4 - j4.9$). Now, by adding a normalized value of $\omega C_1' = 4.9$, we arrive back at a purely resistive case and a 4/1 impedance step-down. The difference between the reactance at point b and at point c is 0.423, which is, of course, the normalized value of $\omega L_2/R_2$.

If instead of a 4/1 impedance step-down we had wanted a 20/1 impedance step-up, we would have continued from point c to the intersection of the 0.1 resistive line and the 0.05 "conductive load circle." This intersection occurs at $0.1 + j1.41$ (point f).

which reflects into $0.05 - j0.705$ (point g); hence adding $\omega C_1 = 0.705 G_2$ will yield the desired 20/1 impedance step-up (point h).

Several things are apparent at this point. One is that component losses may be easily handled since, once normalized, they merely add into the appropriate section of the operation. Another fact is that one is able to work in either direction on the chart; thus, for example, one can assume both C_1 and C_2 and read off the required normalized value of L to tie them together.

To relate Table 9.8–4 to the chart, it is helpful to note that the values of the table correspond to the case in which the path from the equivalent of point h just bisects a step-down reflection circle in the upper right-hand quadrant. Reflecting this bisection point back into the lower left-hand quadrant indicates that either a larger or a small step-down ratio will lead to a smaller value of C_1 and hence to a larger value of $N = C_2/C_1$.

The chart also makes it clear that as Q_2 is increased the variation in the input impedance with frequency or with temperature-induced variations in L and/or the C's will increase.

In addition, the modified chart may be used to find the effect of variations of frequency on Y_{in}. Once a solution is obtained, a 10% increase in the three reactive terms corresponds to a 10% increase in frequency; hence the normalized result is the input admittance at the new frequency. Normally several points suffice to show the pattern of both the real and the imaginary part of Y_{in} near resonance.

Once reasonable values for L, C_1, and C_2 have been located, then one can return to Eq. (9.3–9) or Eq. (9.3–10) and analyze the behavior of the pi network in more detail.

The modified Smith chart is well adapted to dealing with added components such as a capacitor in series with L. (This capacitor might provide dc blocking and/or might allow L to be constructed in a more convenient size. Both of these uses are important at high frequencies.) It is also perfectly capable of handling a two-section filter. Such a filter will have five poles in its transfer impedance and, in addition to the poles, will have four complex zeros in its input impedance. In this case it is even more important that one operate close to the highest-frequency pole or zero so as to avoid poles at the harmonic frequencies. Because the chart is a relatively fast way of obtaining data, one can use it to check the variations in Z_{in} vs. ω to make sure that the circuit is operating near the uppermost pole or zero.†

Pole-Zero Plots

Another way to gain some insight into the passband properties of the pi network is to plot the pole-zero patterns of Z_{in} (and thus the pole patterns of Z_{21}) for several different cases. It must be pointed out that this method is not a satisfactory way to make an initial design, since to make the plots one must first choose the component values rather than being able to work backward from a desirable pattern to component values.

---

† With the middle C equal to twice the outside ones and with equal inductances, such a two-section filter has unloaded poles in its Z_{in} at the origin, at $\pm j\omega_0$, and at $\pm j\sqrt{2}\omega_0$, unloaded zeros in its Z_{in} at $\pm 0.54\omega_0$ and at $\pm j1.31\omega_0$, and $\omega_0^2 = 1/LC_{\text{outside}}$.

As an example consider the two cases in which $N = 1$ ($C_1 = C_2$) and in which $Q_2 = 3$. As we saw in the previous numerical example, there are two different values of ω_0 at which $Z_{in}(j\omega_0)$ is purely resistive. At one frequency $\omega_0 = \omega_2$, while at the other $\omega_0 = 1.342\omega_2$. The first case corresponds to the 9/1 impedance step-down case, while the second corresponds to the 1/1 impedance transformation. In the first case the zeros of $Z_{in}(p)$ occur at $(-0.1667 \pm j0.986)\omega_0$, while in the second case they occur at $(-0.1667 \pm j0.726)\omega_0$. In both cases the negative real pole occurs quite close to $-0.17\omega_0$, while in the $\omega_0 = \omega_2$ case the complex poles lie near $(-0.082 \pm j1.402)\omega_0$ and in the second case they lie near $(-0.0815 \pm j1.042)\omega_0$.

Since in both cases the center frequency is ω_0, a sketch of the pole-zero pattern will indicate that neither $Z_{in}(p)$ nor $Z_{21}(p)$ is too symmetrical around ω_0 for either the 9/1 step-down or for the 1/1 case.

As N is reduced (for a given value of Q_2), the upper complex pole and zero approach each other until at the value of Table 9.8–4 there is only a single frequency at which Z_{in} can be real. For $Q_2 = 3$ the minimum value of N is 0.6. For this value the complex zeros lie at $(0.1667 \pm j0.8975)\omega_0$, while the real pole is near $-0.1275\omega_0$ and the complex poles are near $(-0.1025 \pm j1.1375)\omega_0$. In this case the passband is even more nonsymmetrical than in the $N = 1$ case.

9.9 HIGH-LEVEL AMPLITUDE MODULATION

Single-sideband and suppressed carrier operations are normally done at low signal levels and then amplified up to the desired final power levels in Class B "linear" RF amplifiers.

While normal amplitude modulation can be done at lower power levels, it is usually accomplished as close to the final stage as possible. What we shall try to do in this section is explain the ideal goal of all high-level modulating circuits and how one goes about achieving this goal. We shall not be able to provide specific details about particular designs. It is hoped that, if the reader knows what he is trying to do, it will be easier for him to do it properly.

Before one tries to modulate the signal, it is wise to compare the overall efficiencies of Class A and Class B RF stages when each is amplifying a 100% sinusoidally modulated signal. The results are as follows:

$$\begin{aligned} \eta &= \frac{3}{8}\,\frac{1}{2}\,\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \qquad \text{for Class A,} \\ \eta &= \frac{3}{4}\,\frac{\pi}{4}\,\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \qquad \text{for Class B.} \end{aligned} \tag{9.9–1}$$

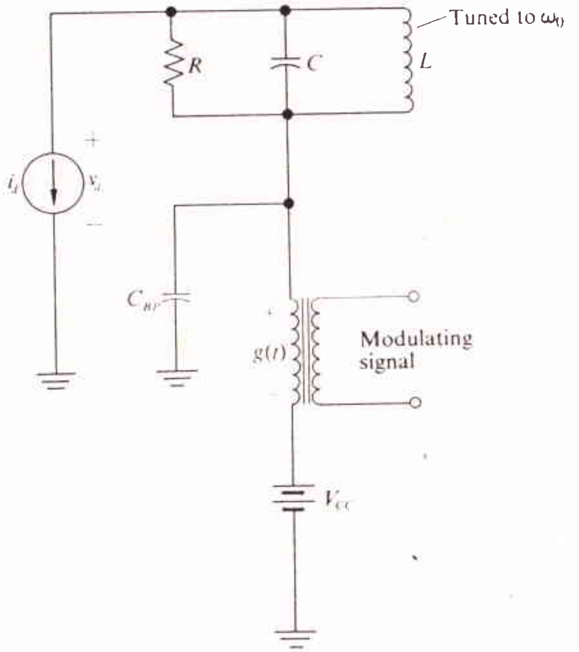
With the same carrier but no modulation, the efficiencies drop to

$$\begin{aligned} \eta &= \frac{1}{8}\,\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \qquad \text{for Class A,} \\ \eta &= \frac{\pi}{8}\,\frac{V_{CC} - V_{\text{sat}}}{V_{CC}} \qquad \text{for Class B.} \end{aligned} \tag{9.9–2}$$

Thus for normal variable amplitude modulation the efficiency of the Class B amplifier lies between 50% and 75% of its peak efficiency, while the Class A stage achieves only 25% to 37.5% of its peak efficiency. For either no modulation or full modulation, the Class B efficiency is π times the Class A efficiency.

A logical first step toward an ideal high-level modulator would be to try to raise the average efficiency of the Class B amplifier from the 35–55% range found here back to the 75% possible for unmodulated signals.

Figure 9.9–1 illustrates a circuit that allows us to have linear envelope amplification at a constant theoretical efficiency of nearly $\pi/4$. In this circuit C_{BP} is presumed to be a short circuit with respect to RF frequencies and an open circuit with respect to the "modulating" signal $f(t)$.



$$i_d = [1 + f(t)]\left[\frac{I_p}{\pi} + \frac{I_p}{2}\cos\omega_0 t + \cdots\right]$$

$$g(t) = f(t)\left[V_{CC} - V_{\text{sat}}\right]$$

Fig. 9.9–1 Output-modulated Class B amplifier for a modulated signal.

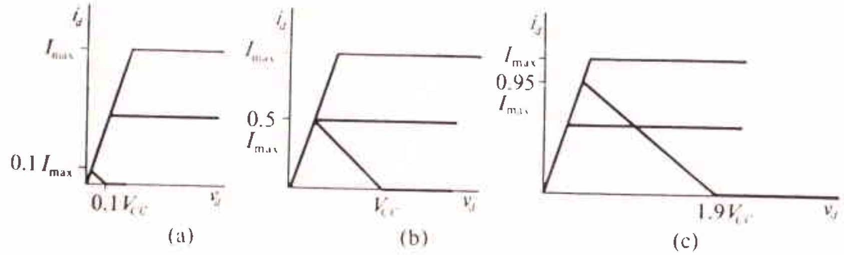
Figure 9.9–2 shows the operating paths for values of $f(t)$ of -0.9, 0.0, and $+0.9$ for a device with a linear saturation characteristic that passes through the origin. For this device characteristic (which is a good approximation for many transistors), the circuit is always operating as a fully driven Class B amplifier; hence its efficiency is always

$$\eta = \frac{\pi}{4}\left(1 - \frac{I_p R_{\text{sat}}}{2V_{CC}}\right), \tag{9.9–3}$$

where R_{sat} is the slope of the device saturation characteristic and I_p is the peak current at 100% modulation. Since peak ac power output for this case is $(2V_{CC} - I_p R_{\text{sat}})(I_p/4)$, the ratio of peak ac power output to peak device dissipation becomes even larger than the Class B result of Eq. (9.2–3) for unmodulated carriers. When $I_p R_s = V_{\text{sat}} = (0.05)2V_{CC}$, then

$$\eta = 0.95\frac{\pi}{4}, \qquad P_{\text{ac}} = 0.95\frac{V_{CC}I_p}{2}, \qquad P_{\text{dc}} = \frac{2V_{CC}I_p}{\pi},$$
$$P_{\text{device}} = 0.162\,V_{CC}I_p, \qquad \text{and} \qquad \frac{P_{\text{ac}}}{P_{\text{device}}} = 2.95. \tag{9.9–4}$$

Hence for a 100 W output on a modulation peak or a 25 W unmodulated output all one needs is a device capable of dissipating a peak power of 100/2.95 = 34 W. If the period of the lowest-frequency modulating signal is sufficiently shorter than the thermal time constant of the device so that one can assume averaging over a modulation cycle, then for a 100% sinusoidally modulated signal a 12.8 W device dissipation capability should be sufficient. (This is in fact the normal situation.) Thus, by "modulating" the power supply of an amplifier driven by a modulated signal, we have increased the long-term average efficiency by a factor of about 2/1 and reduced the device's power-handling requirements by more than 3/1.



(a) (b) (c)

Fig. 9.9–2 Operating paths for a modulated Class B amplifier, "variable supply."

Note that, for an unmodulated input to the characteristic of Fig. 9.9–2(b), increasing $V_{CC}[1 + f(t)]$ upward from V_{CC} would not lead to any variation in the output whatsoever. Reduction of $V_{CC}[1 + f(t)]$ downward from V_{CC} would lead to the combination of effects discussed in Section 8.6.

As was explained in Chapter 8, with the drive properly increased, the collector modulation and its accompanying "reflection" into the base circuit can lead to a linear modulator.

All in all, the desirable operation of a transistorized AM transmitter would seem to require that the output stage be of the type illustrated in Figs. 9.9–1 and 9.9–2. For high-power outputs one would extend this reasoning back to include the driver stage. At some point where the driver power level is reasonably low, one must generate the AM either by a collector saturation technique or possibly by one of the other methods outlined in Chapter 8.

One further method that can be used to improve the quality of a modulation system is to employ one of the AM detectors from Chapter 10 to detect the final

9.9

modulation and to apply it via a negative feedback circuit to correct the minor distortions of the whole modulation chain.

Modulated Vacuum-Tube Amplifiers

If a modulation transformer were included in series with V_{BB} in Fig. 9.4–1 (and if it was adequately bypassed for RF signals), then the effect of applying slow modulation would be the same as the effect of varying V_{BB}. Actually, if we wished to modulate this amplifier, we would first convert it to the form of Fig. 9.8–1. Furthermore, we would choose the time constant of R_G and C_G so that the bias circuit would be able to track the highest modulation frequency. With these changes, as well as a reduction in the drive signal and a slight reduction in the load impedance, we would have a modulated Class C amplifier.

It is not at all obvious intuitively why such a modulator should be linear. It is a fact that reasonable linearity can be obtained through the judicious combination of three effects:

a) The finite r_p of the tube makes i_B a function of v_{PK}.
b) The grid current is also a function of v_{PK}; hence the grid circuit loading, and therefore the drive voltage, is a function of the plate-cathode voltage.
c) The grid circuit bias is a function of I_{G0} and hence of v_{PK}.

Calculation of these effects is complicated by the fact that now the tank load impedance must be fixed, it can no longer be left free to be defined as V_1/I_1. Thus one is forced to assume a tank voltage, a drive voltage, and a bias for a given combination of V_{BB} and modulation and to derive an operating path and set of current pulses. These current pulses must now yield values for I_{G1}, I_1, and I_{G0} that are consistent with the assumed values. Obviously with so many variables this is not a simple procedure.

Intuitively what one wants to happen is this: As the modulation drops into a valley, the bias should increase, and the loading should reduce the drive so that the output swing falls but the efficiency holds constant. Then as the modulation swings toward a peak, the drive should increase, and the bias should fall so that again the efficiency remains reasonably constant.

In effect, we are trying to reflect the modulation into the grid circuit. The fact that this "reflection" combines with the nonlinearity in the relationship between the current pulse size and shape and its fundamental value to yield an overall linear result is just a fortuitous circumstance that we exploit.

Many commercial high-power modulations are claimed to have plate efficiencies up to 80%. Since this is only 5% better than is possible with the "variable collector voltage" Class B modulated amplifier, the vacuum tube has very little theoretical advantage as a high-level modulator. Actual overall tube efficiencies are always lower than transistor efficiencies, since the vacuum tube requires an additional power source of 10% to 20% of its available output power just to heat its filaments.

Most higher-level transmitters currently available employ vacuum tubes because of the tubes' high voltage ratings and hence their ability to produce high power at reasonable current levels rather than because of any theoretical superiority in conversion efficiency.

PROBLEMS

9.1 In the circuit of Fig. 9.P–1, $i_i(t) = (0.1 + 0.09 \cos \omega t)$ A and the transistor β is constant at 20 for $4\text{ A} > i_C > 100\text{ mA}$ (this is unlikely in practice). Find the power delivered by the battery, the power consumed in the 3 Ω resistor, and the power dissipated in the transistor.

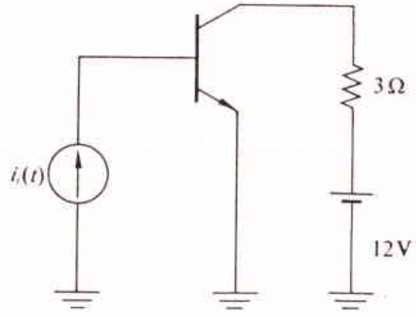


Figure 9.P–1

9.2 Assuming that the transistor in Problem 9.1 has a thermal resistance of 4°C/W from its junction to its case, find the temperature rise of the junction over the case both for the signal of Problem 9.1 and for the case where the ac portion of the drive signal is removed.

9.3 Assume that the 3 Ω resistor of Problem 9.P–1 is replaced by the primary of a lossless transformer that reflects a 3 Ω load at the operating frequency ω. Repeat Problems 9.1 and 9.2 for this case. Make the additional assumption that the 12 V supply is reduced to 6 V.

9.4 Repeat Problem 9.3 assuming that the 0.09 peak amplitude sine-wave drive is replaced by a 0.09 A peak amplitude square wave.

9.5 Compare the conversion efficiencies of the cases of Problems 9.1, 9.3, and 9.4. Explain physically the great difference in these values.

9.6 Repeat Problem 9.4 assuming that the broadband transformer-coupled load is replaced by a narrowband tuned circuit that presents the same impedance at the fundamental frequency.

9.7 In Problem 9.4 assume that the square-wave drive is not constant at 90 mA but that its amplitude varies so that the percentage of time at each amplitude is as listed in Table 9.P–1. Find the long-term average values of collector dissipation, load power, and power supplied by the battery.

Table 9.P–1

Percent of time	Peak amplitude, mA
10	90
20	80
30	70
20	50
10	30
10	10

9.8 Find the "long-term average" power supplied to a 50 Ω load by each of the following signals. ("Long-term average" means averaged over a complete cycle of the longest period in the modulation.)

a) A triangular modulated 10 kHz sine wave as shown in Fig. 9.P–2(a).
b) A repetitive wave as shown in Fig. 9.P–2(b).
c) A 50 V rms, 5 MHz carrier that is frequency-modulated to an rms deviation of 50 kHz by Gaussian noise.
d) A square-wave-and-generator combination as shown in Fig. 9.P–2(c).
e) A tone burst controlled triangular signal as shown in Fig. 9.P–2(d).

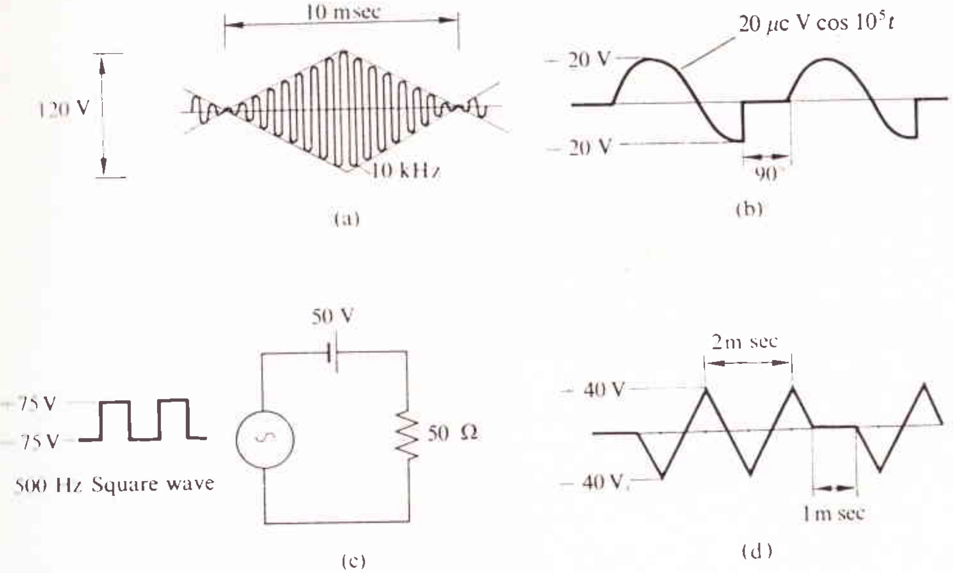


Figure 9.P–2

9.9 A certain device produces an output current that flows in half-sine-wave pulses of peak amplitude equal to 2.88 A and peak spacing of 1 μsec. The minimum allowable output voltage of the device without excessive distortion is 1 V. Choose values for C_1, C_2, and L (assume L is lossless) that will yield maximum power to the 72 Ω load while preventing excessive distortion across C_1 of Fig. 9.P–3 $\omega_0 C_2 = 0.07$ mho. Calculate the size of the second harmonic component of the voltage across both C_1 and across the 72-Ω load. Sketch the passband for ±10% around the center frequency for transmission from the device current pulses to the 72-Ω load.

9.10 Consider the circuit shown in Fig. 9.3–1 with a 120° wide pulse train wherein $I_p = 2$ A. If the resonant impedance of the tuned circuit is 3 kΩ and the allowable minimum device voltage is 200 V, what value must V_{CC} have? Find the fundamental output power for this case. If the tuned-circuit Q is 10, calculate the magnitude and phase of the output second-harmonic voltage. What is the conversion efficiency of this device under these conditions? What is the device dissipation?

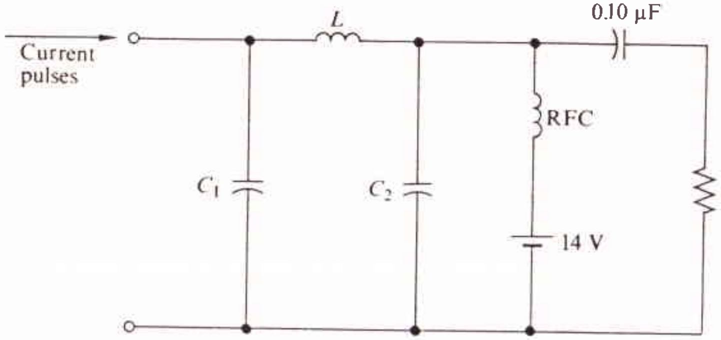


Figure 9.P–3

9.11 The circuits of Fig. 9.5–3(a) and 9.6–3(c) are both driven by a 10 kHz, 1 V peak square wave. The positive supply in each case is 20 V, while the load resistor is 10 Ω ($R_E = 0.25$ Ω and the coupling capacitor is 100 μF in Fig. 9.6–3c). Draw the transistor collector current waveshapes for both cases and calculate the power delivered to the 10 Ω load. [The series LC in Fig. 9.5–3(a) is resonant at 10 kHz.] Define or specify any other parameters as necessary to complete the problem.

9.12 In the circuit of Fig. 9.6–3(a), $V_{CC} = 20$ V, $\beta \approx 50$, $R_L = 100$ Ω, $R_1 = 25$ kΩ, $R_2 = 1$ kΩ, and the input voltage is successively a train of 13.5 V peak sine, square, and triangular waves, each with zero average value. Calculate the load power and the transistor dissipation for each case. Make, and state, any assumptions that are necessary in working the problem.

9.13 For the circuit of Problem 9.12, find the amplitude of each driving waveshape that results in maximum transistor dissipation. Assume a zero signal emitter current of 2 mA per transistor.

PULSE TRAIN EXPANSIONS

In general it is possible to approximate the plate, grid, collector, or base current waveshapes that are produced in Class C amplifiers by one of several tabulated waveshapes (or by some combination of these waveshapes).

To aid in such calculations, this appendix tabulates the general results of I_n/I_p vs. τ/T for six waveshapes. In addition, it presents graphs of I_{dc}/I_p, I_1/I_p, and I_2/I_p all versus τ/T or the total conduction angle, θ.

Figures 4.2-3 and 4.2-4 present similar data for sine-wave tip curves. In practice, the sine-wave tip and the cosine pulse data are almost interchangeable in the region where $\tau/T \leq \frac{1}{2}$.

Table 9.A-1

Pulse shape	I_{dc}	I_n
Rectangular	$\frac{I_p\tau}{T}$	$\frac{2I_p}{n\pi}\sin\frac{n\pi\tau}{T}$
Triangular	$\frac{I_p\tau}{2T}$	$\frac{I_p\tau}{T}\frac{\sin(n\pi\tau/2T)^2}{n\pi\tau/2T}$
Symmetrical trapezoidal	$\frac{3I_p\tau}{4T}$	$\frac{3I_p\tau}{2T}\frac{\sin(n\pi\tau/4T)}{n\pi\tau/4T}\frac{\sin(3n\pi\tau/4T)}{3n\pi\tau/4T}$
Cosine pulse	$\frac{2I_p}{\pi}\frac{\tau}{T}$	$\frac{4}{\pi}I_p\frac{\tau}{T}\frac{\cos(n\pi\tau/T)}{1-(2n\tau/T)^2}$
Cosine-squared pulse	$\frac{I_p\tau}{4T}$	$\frac{I_p}{n\pi}\frac{\sin(n\pi\tau/2T)}{1-(n\tau/2T)^2}$

The cosine pulse is not a sine-wave tip (except when $\tau/T = \frac{1}{2}$), although it closely approximates one; $i = I_p\cos(\pi t/\tau)$ if $|t| < \tau$, $i = 0$ otherwise. The cosine-squared pulse is equal to $I_p\cos^2(\pi t/\tau)$ when $|t| < \tau/2$ and zero otherwise. At $t = \pm\tau/4$, this pulse has a value of $I_p/2$ while the cosine pulse has a value of $I_p/\sqrt{2}$. Table 9.A-1 and Fig. 9.A-1 show the waveshapes and the expressions for I_{dc} and I_n. Figures 9.A-2, 9.A-3, and 9.A-4 show the normalized fundamental, dc, and second-harmonic components vs. the conduction angle.

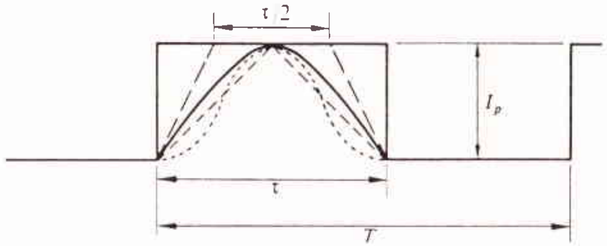


Fig. 9.A-1 Waveshapes illustrated in Table 9.A-1.

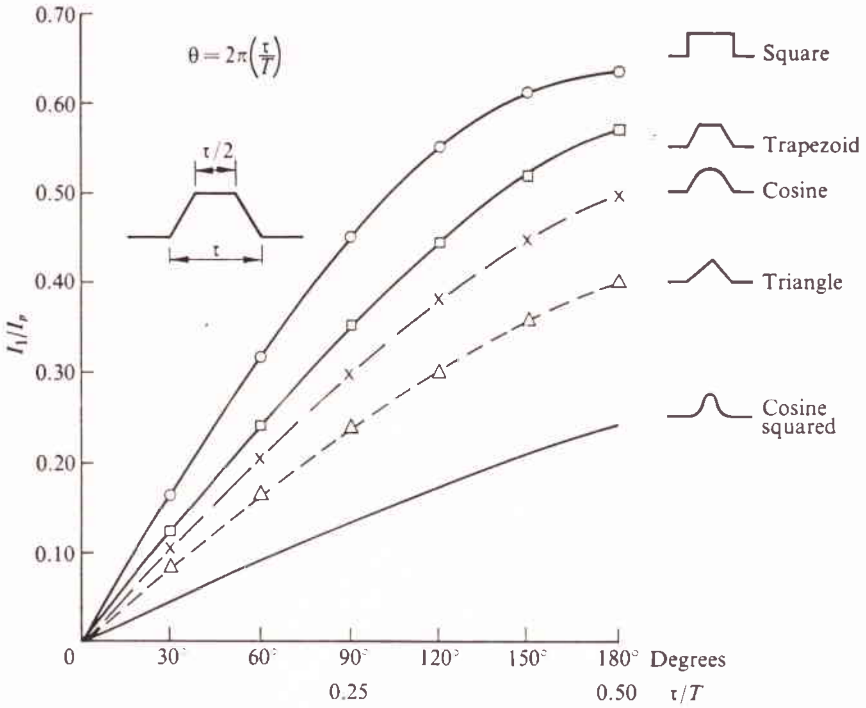


Fig. 9.A-2 I_1/I_p vs. τ/T or total conduction angle for repetitive pulse trains.

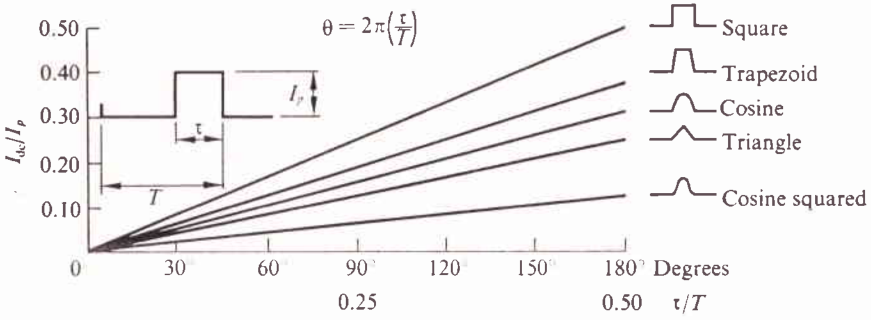


Fig. 9.A-3 I_{dc}/I_p vs. τ/T or total conduction angle for repetitive pulse trains.

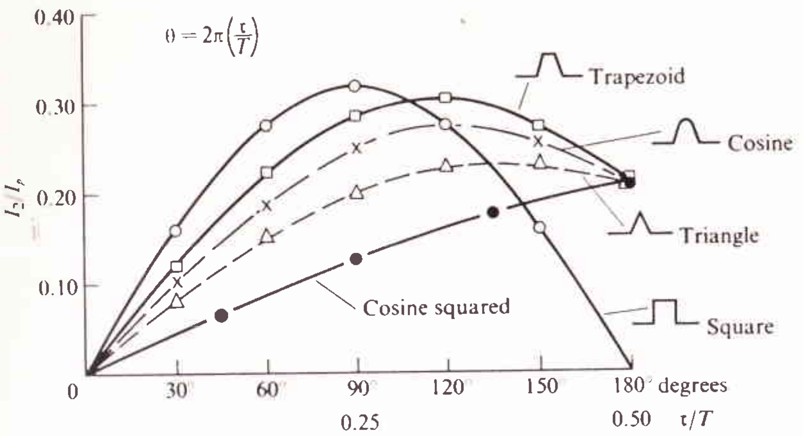


Fig. 9.A–4 I_2/I_p vs. τ/T or total conduction angle in degrees for repetitive pulse trains.

CHAPTER 10

AMPLITUDE DEMODULATORS

In this chapter we consider the theoretical limitations on the amplitude demodulation of signals, as well as the details of a number of practical circuits to accomplish this demodulation.

The most widely used practical circuit is the narrowband envelope detector of Sections 10.3 and 10.4. Though simple to construct, this circuit has definite limitations that may be circumvented by the use of synchronous or average detectors. In other sections of the chapter we examine these circuits in some detail.

10.1 AMPLITUDE DEMODULATION TECHNIQUES

In this section we consider the basic theoretical principles involved in the demodulation of AM and SSB signals. In the subsequent sections of this chapter we shall consider the practical circuits by which these principles are implemented. In general, there are only three basic methods of amplitude demodulation or detection: (a) synchronous detection, (b) average envelope detection, and (c) peak envelope detection. The basic idea behind all of these methods is to reclaim the modulation information $g(t)$ from the modulated carrier which has the form

$$v_i(t) = g(t)\cos\omega_0 t, \qquad \begin{cases} \text{normal AM:} & g(t) \geq 0, \\ \text{suppressed carrier AM:} & \overline{g(t)} = 0, \end{cases}$$

$$v_i(t) = \frac{g(t)}{2}\cos\omega_0 t \pm \frac{\hat{g}(t)}{2}\sin\omega_0 t, \qquad \text{(SSB)} \tag{10.1–1}$$

where $\hat{g}(t)$ is the Hilbert transform of $g(t)$ (see Chapter 8). As we shall see, synchronous detection must always be employed to demodulate SSB or suppressed carrier AM, while any of the three demodulation methods may be employed to demodulate a normal AM signal whose modulation index does not exceed unity.

Synchronous Detection

The block diagram of the synchronous detector is shown in Fig. 10.1–1. For this detector, if $v_i(t)$ has the form given by Eq. (10.1–1), then with the aid of the identities

$$\cos x \cos y = \tfrac{1}{2}[\cos(x + y) + \cos(x - y)]$$

and

$$\sin x \cos y = \tfrac{1}{2}[\sin(x + y) + \sin(x - y)]$$

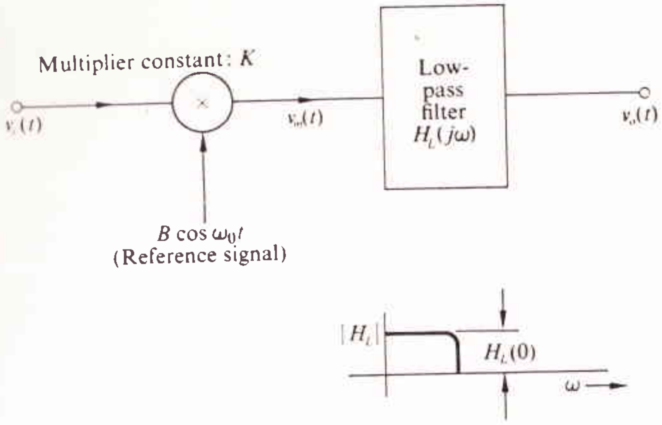


Fig. 10.1–1 Block diagram of synchronous detector.

he multiplier output v_m can be written as

$$v_m(t) = KB\left[\frac{g(t)}{2} + \frac{g(t)}{2}\cos 2\omega_0 t\right], \qquad \text{(normal or suppressed carrier AM)}$$

$$v_m(t) = KB\left[\frac{g(t)}{4} + \frac{g(t)}{4}\cos 2\omega_0 t \pm \frac{\hat{g}(t)}{4}\sin 2\omega_0 t\right]. \qquad \text{(SSB)} \tag{10.1–2}$$

If the low-pass filter removes the components of $v_m(t)$ which are concentrated about the radian frequency $2\omega_0$, then the detector output takes the form

$$v_o(t) = \frac{KBg(t)}{\lambda} * h_L(t), \tag{10.1–3}$$

where $\lambda = 2$ for normal AM, $\lambda = 4$ for SSB, and $h_L(t)$ is the impulse response of the low-pass filter. If the filter has a wide enough bandwidth to pass $g(t)$ undistorted, then Eq. (10.1–3) simplifies to the desired form

$$v_o(t) = \frac{KBg(t)}{\lambda}H_L(0), \tag{10.1–4}$$

where $H_L(j\omega)$ is the Fourier transform of $h_L(t)$. Note that if the reference signal has the form $B\cos(\omega_0 t + \theta)$, then $v_0(t)$ is given by (for normal AM)

$$v_o(t) = \frac{KBg(t)H_L(0)\cos\theta}{2},$$

which is an attenuated version of Eq. (10.1–4). Hence for maximum detector output, θ must be designed to equal zero.

For synchronous detection to be accomplished it is apparent that the spectral components of $v_m(t)$ in the vicinity of $2\omega_0$ must not overlap the spectral components of $v_m(t)$ in the vicinity of the origin. For normal or suppressed carrier AM the Fourier transform $V_m(\omega)$ of $v_m(t)$ is given by

$$V_m(\omega) = \frac{KBG(\omega)}{2} + \frac{KB}{4}[G(\omega - 2\omega_0) + G(\omega + 2\omega_0)], \tag{10.1–5}$$

where $G(\omega)$ is the Fourier transform of $g(t)$. A plot of $|V_m(\omega)|$ vs. ω for the case where $g(t)$ is band-limited in frequency to ω_m is shown in Fig. 10.1–2. In order to separate the desired output signal from the double frequency term, the inequality

$$\omega_m < \omega_0$$

must be satisfied; i.e., the maximum modulation rate must be less than the carrier frequency. If ω_0 is not greater than ω_m, then synchronous detection (or, for that matter, any other form of detection) is impossible. In addition, the closer ω_m is to ω_0, the more complex the low-pass filter of the synchronous detector must be to extract the desired output signal.

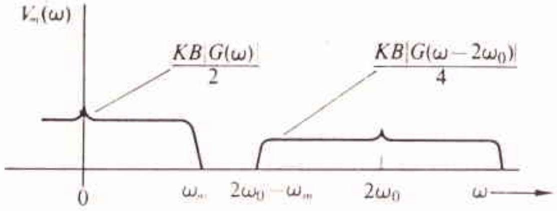


Fig. 10.1–2 Spectrum of $|V_m(\omega)|$ vs. ω with $g(t)$ band-limiting to ω_m.

For an SSB signal with $g(t)$ bandlimited to ω_m, $|V_m(\omega)|$ vs. ω has the form shown in Fig. 10.1–3. Here we see that for lower-sideband SSB, synchronous detection is possible only if $\omega_m < \omega_0$, whereas for upper-sideband SSB synchronous detection can be accomplished only if $\omega_m < 2\omega_0$.

To implement the synchronous detector, a multiplier and a reference-signal source of the form $B\cos\omega_0 t$ are required. Any of the multiplier circuits discussed in Chapter 7 or 8 is sufficient as a multiplier. After all, synchronous detection is essentially mixing down to dc. The reference-signal source at ω_0 may be obtained by placing $v_i(t)$ through a very narrowband filter which extracts the carrier component of $v_i(t)$ (if it exists) and removes the sideband information. Specifically, if

$$v_i(t) = A[1 + mf(t)]\cos\omega_0 t,$$

where the Fourier transform $V_i(\omega)$ of $v_i(t)$ and the spectrum $F(\omega)$ of $f(t)$ are as shown in Fig. 10.1–4, then the output of the narrowband filter centered at ω_0 with a bandwidth less than $2\omega_L$ is given by

$$A\cos\omega_0 t,$$

which is the desired reference signal. Such an extremely narrowband filter is usually implemented with a crystal or a mechanical filter or with a phase-locked loop (PLL). If a crystal filter is employed, the center frequency ω_0 of $v_i(t)$ must also be crystal

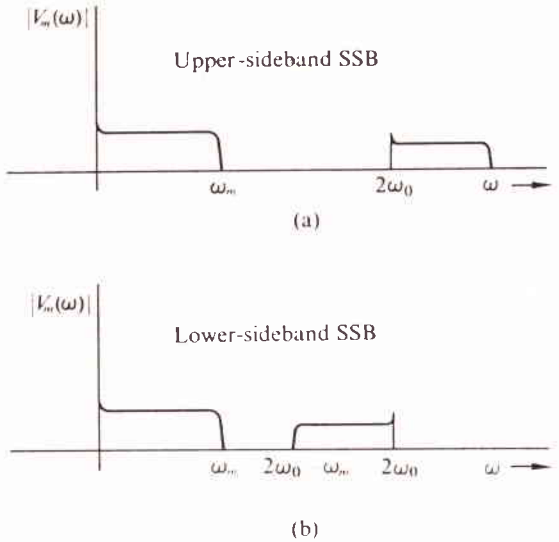


Fig. 10.1–3 Plot of $|V_m(\omega)|$ vs. ω for SSB modulation with $g(t)$ band-limited to ω_m.

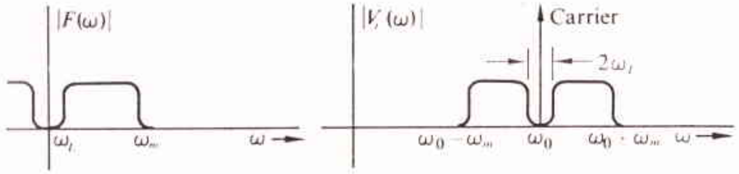


Fig. 10.1–4 Plot of $|F(\omega)|$ and $|V_i(\omega)|$ vs. ω.

controlled. On the other hand, if a PLL is employed, ω_0 need not be (but usually is) crystal controlled, since the phase-locked loop is capable of tracking slow variations in ω_0 while still acting as a narrowband filter about ω_0.

In the case of SSB and suppressed carrier AM, a carrier component does not usually exist in $v_i(t)$. For SSB signals a low-level or pilot carrier is usually added to the transmitted signal and extracted as the reference signal at the receiver with a narrowband filter.† However, when the SSB modulation is an audio signal, the reference signal may be obtained from an independent oscillator. This oscillator generates a signal of the form $B \cos [\omega_0 t + \theta(t)]$, where $\theta(t)$ is a slowly varying random phase (or frequency) error between the oscillator at the transmitter and the oscillator at the receiver. With the independent reference signal, the output of the synchronous detector is given by

$$v_o(t) = \frac{KB}{4}[g(t) \cos \theta(t) \pm \hat{g}(t) \sin \theta(t)]. \tag{10.1–6}$$

Since it has been observed that the human ear cannot readily distinguish between $g(t)$ and $\hat{g}(t)$, the output signal of the detector is a perfectly acceptable reproduction of $g(t)$ even when $\theta(t)$ drifts through $\pi/2$. If $\theta(t) = \varepsilon t + \theta$, that is, if the reference oscillator differs in frequency from the transmitter oscillator, the effect on the ear is to raise or lower the pitch of the demodulated signal while maintaining its intelligibility. The eye, however, is not as forgiving as the ear; hence an independent reference oscillator cannot be used for demodulating SSB video information.

For suppressed carrier AM the reference signal can be generated by passing $v_i(t)$ through a network of the form shown in Fig. 10.1–5. The output of the square-law device, which may be implemented with a field effect transistor, is given by

$$\begin{aligned} v_2(t) &= kg^2(t) \cos^2 \omega_0 t \\ &= \frac{kg^2(t)}{2} + \frac{kg^2(t)}{2} \cos 2\omega_0 t. \end{aligned} \tag{10.1–7}$$

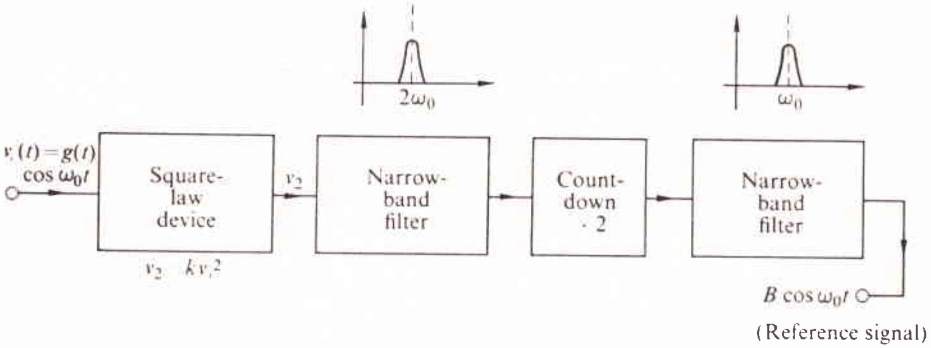


Fig. 10.1–5 Network for extracting reference carrier from suppressed-carrier AM.

Since $g^2(t)$ is always greater than zero, its average value, unlike the average value of $g(t)$, is greater than zero and thus $v_2(t)$ has a carrier frequency component at $2\omega_0$ which is in turn extracted by the narrowband filter centered at $2\omega_0$. This filter may also be implemented as a crystal filter or as a phase-locked loop. The filter output is then counted down in frequency by 2 and filtered to provide the desired reference signal $B \cos \omega_0 t$. The coefficient B is, of course, a function of the individual device and filter scale factors within the network of Fig. 10.1–5.

Average Envelope Detection

The block diagram of the average envelope detector is shown in Fig. 10.1–6. If $v_1(t)$ is a normal AM input of the form

$$v_1(t) = g(t) \cos \omega_0 t,$$

---

† In many voice communication systems a large number of voice channels ($BW < 3$ kHz) are SSB-modulated and frequency-division-multiplexed at 4 kHz intervals from dc to $4n$ kHz, where n is the total number of channels. This composite signal plus a 4 kHz reference signal then frequency-modulates a high-frequency carrier. At the receiver, after frequency demodulation the various SSB signals must be synchronously demodulated. The approximate reference signal for each SSB channel is generated by extracting the 4 kHz reference signal and then distorting it and extracting the desired harmonic by appropriate filtering.

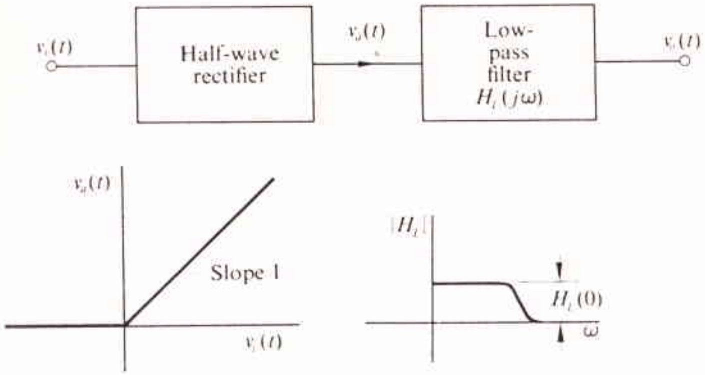


Fig. 10.1–6 Block diagram of average envelope detector.

where $g(t) \geq 0$, and if the transfer characteristic of the half-wave rectifier is given by

$$v_a = \begin{cases} v_i, & v_i > 0, \\ 0, & v_i \leq 0, \end{cases} \tag{10.1–8}$$

as shown in Fig. 10.1-6, then

$$v_a(t) = g(t)(\cos \omega_0 t)S(t), \tag{10.1–9}$$

where $S(t)$ is a switching function with the property $S(t) = 1$ for $\cos \omega_0 t > 0$ and $S(t) = 0$ for $\cos \omega_0 t \leq 0$. A sketch of $v_i(t)$, $v_a(t)$, and $S(t)$ is shown in Fig. 10.1–7.

It is intuitively obvious from the sketch that by extracting the average value of $v_a(t)$ with the low-pass filter, an output signal proportional to $g(t)$ is obtained. To demonstrate rigorously that $v_o(t)$ is proportional to $g(t)$, we expand $S(t)$ in its Fourier series

$$S(t) = \frac{1}{2} + \frac{2}{\pi}\cos \omega_0 t - \frac{2}{3\pi}\cos 3\omega_0 t + \cdots \tag{10.1–10}$$

and substitute it into Eq. (10.1–9) to obtain

$$\begin{aligned} v_a(t) &= \frac{g(t)}{2}\cos \omega_0 t + \frac{2g(t)}{\pi}\cos^2 \omega_0 t - \frac{2g(t)}{3\pi}\cos 3\omega_0 t \cos \omega_0 t \\ &= \frac{g(t)}{\pi} + \frac{g(t)}{2}\cos \omega_0 t + \text{higher harmonic AM signals.} \end{aligned} \tag{10.1–11}$$

If the low-pass filter (cf. Fig. 10.1–6) removes the frequency components of $v_a(t)$ centered about ω_0, $2\omega_0$, etc., then $v_o(t)$ is given by

$$v_o(t) = \frac{g(t)}{\pi} * h_L(t), \tag{10.1–12}$$

where $h_L(t)$ is the impulse response of the low-pass filter. If, in addition, the low-pass filter has a wide enough bandwidth to pass $g(t)$ undistorted, then Eq. (10.1–12) simplifies to the desired form

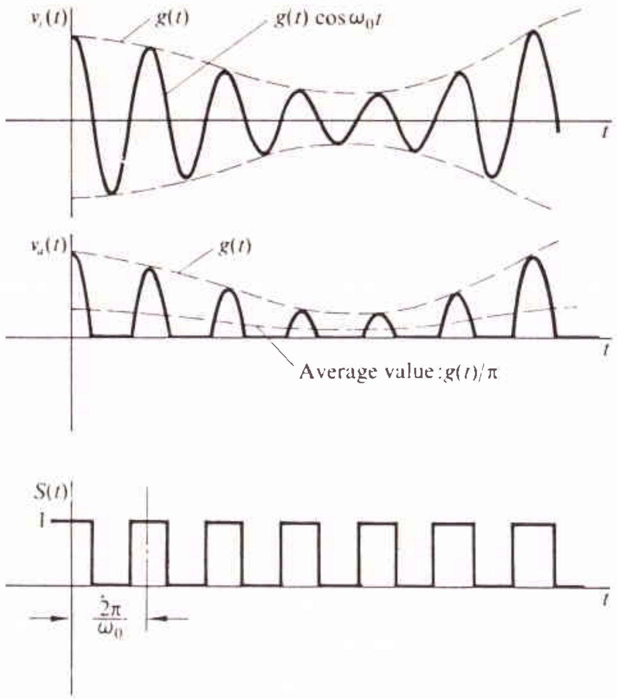


Fig. 10.1–7 Plot of $v_i(t)$, $v_a(t)$ and $S(t)$ vs. t.

$$v_o(t) = \frac{g(t)}{\pi}H_L(0), \tag{10.1–13}$$

where again $H_L(j\omega)$ is the Fourier transform of $h_L(t)$.

If we plot the magnitude of the Fourier transform $V_a(\omega)$ of $v_a(t)$ for the case where $g(t)$ is band-limited to ω_m as shown in Fig. 10.1–8, we observe that the inequality

$$\omega_m < \frac{\omega_0}{2} \tag{10.1–14}$$

must be satisfied if the low-pass filter is to be capable of extracting $g(t)$ from $v_a(t)$.

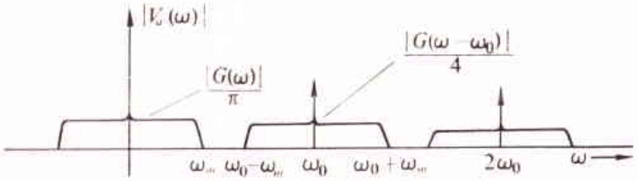


Fig. 10.1–8 Plot of $|V_a(\omega)|$ vs. ω.

If the half-wave rectifier of Fig. 10.1–6 is replaced by a full-wave rectifier, not only does the output of the low-pass filter double, but no term centered about ω_0 exists in $v_a(t)$. Consequently $g(t)$ may be extracted from $v_a(t)$, provided that $\omega_m < \omega_0$, which is the same frequency constraint imposed by the synchronous detector.

It is apparent from Fig. 10.1–7 that the output signal $v_o(t)$ must always be positive,† i.e., the average envelope detector extracts the positive envelope. If $g(t) \cos \omega_0 t$ were replaced by $-g(t) \cos \omega_0 t = g(t) \cos (\omega_0 t + \pi)$, the output would be unchanged, since the average envelope detector is clearly insensitive to the phase of the carrier; hence the output of the average envelope detector is written more precisely as

$$v_o(t) = \frac{|g(t)|}{\pi} H_L(0). \tag{10.1–15}$$

Thus if a suppressed carrier wave for which $g(t)$ has both positive and negative values were demodulated by an envelope detector, gross distortion would appear at the output. Specifically, the output would appear as if the modulation had been passed through a full-wave rectifier as illustrated in Fig. 10.1–9. Similarly, if the modulation index of a normal AM signal exceeds unity, the average envelope detector introduces distortion and synchronous detection must be employed.

By writing the expression for the SSB signal given by Eq. (10.1–1) in the equivalent form

$$v_i(t) = \tfrac{1}{2}\sqrt{g^2(t) + \hat{g}^2(t)} \cos \left[\omega_0 t + \tan^{-1} \frac{\hat{g}(t)}{g(t)} \right], \tag{10.1–16}$$

we observe that the output of an average envelope detector with an SSB input would take the form

$$v_o(t) = \frac{1}{2\pi}\sqrt{g^2(t) + \hat{g}^2(t)} H_L(0). \tag{10.1–17}$$

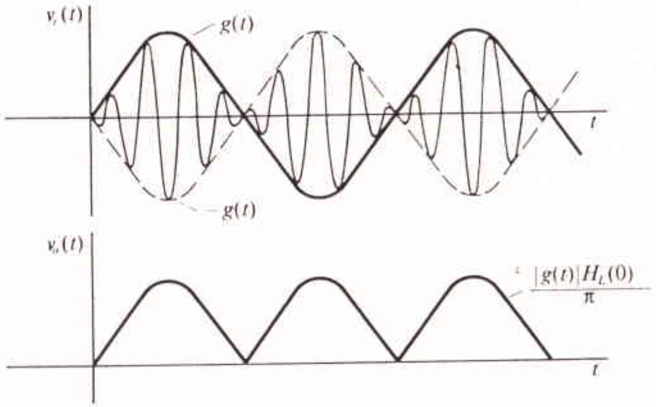


Fig. 10.1–9 Output of average envelope detector.

---

† If the half-wave rectifier characteristics are reversed, that is, if $v_a = v_i$ for $v_i < 0$ and $v_a = 0$ for $v_i \geq 0$, then $v_o(t)$ must always be negative.

which is also a highly distorted output signal. Here again we see that synchronous detection is required.

However, if a large carrier component $A \cos \omega_0 t$ is added to the transmitted SSB signal, the resultant signal may be demodulated by envelope detection. Specifically, if

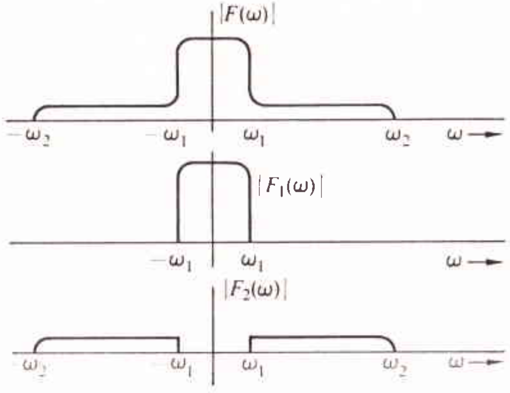
$$\begin{aligned} v_i'(t) &= v_i(t) + A \cos \omega_0 t \\ &= \left[A + \frac{g(t)}{2} \right] \cos \omega_0 t - \frac{\hat{g}(t)}{2} \sin \omega_0 t \\ &= \sqrt{\left[A + \frac{g(t)}{2} \right]^2 + \frac{\hat{g}^2(t)}{4}} \cos \left[\omega_0 t + \tan^{-1} \frac{\hat{g}(t)}{2A + g(t)} \right] \end{aligned} \tag{10.1–18}$$

is demodulated by an envelope detector, the output is proportional to

$$\sqrt{\left[A + \frac{g(t)}{2} \right]^2 + \frac{\hat{g}^2(t)}{4}} \approx \frac{A + g(t)}{2}$$

provided that $A \gg |g(t)/2|$ and $A \gg |\hat{g}(t)/2|$. Such a signal is rarely transmitted, since most of the transmitter power must be expended in providing a large carrier.

The basic concept, however, provides the basis for vestigial sideband transmission. Most video signals have the property that their frequency spectra are concentrated at low frequencies, as shown in Fig. 10.1–10. Consequently, when a carrier at ω_0 is amplitude-modulated by a video signal, the resultant signal $v_1(t)$ has a bandwidth of $2\omega_2$. This bandwidth may be reduced considerably to $\omega_1 + \omega_2$ by placing the AM signal through a filter $H(j\omega)$ of the form shown in Fig. 10.1–11, which removes the low-level portion of the lower- (or upper-) sideband information and doubles the low-level portion of the upper- (or lower-) sideband information to produce a vestigial sideband signal $v_2(t)$. The basic property of the resultant vestigial sideband signal is that it requires less bandwidth for its transmission and still is capable of being demodulated by an envelope detector.



$F(\omega) = F_1(\omega) + F_2(\omega)$

Fig. 10.1–10 Frequency spectrum of typical video signal $f(t)$.

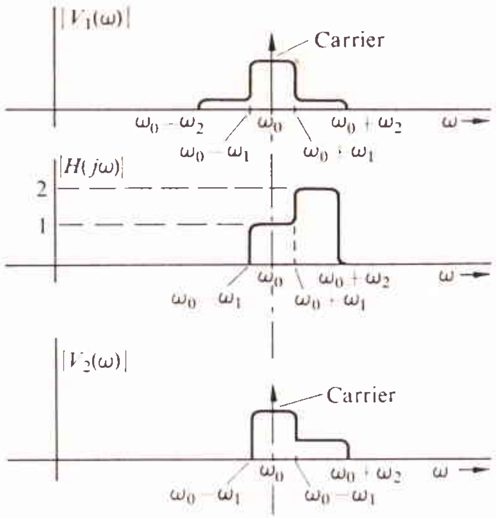


Fig. 10.1–11 Construction of a vestigial sideband signal.

To demonstrate this fact, we write the video signal $f(t)$ in the form

$$f(t) = f_1(t) + f_2(t), \tag{10.1–19}$$

where $f_1(t)$ corresponds to the high-level, low-frequency portion of $f(t)$ and $f_2(t)$ corresponds to the low-level, high-frequency portion of $f(t)$. The resultant AM signal therefore takes the form

$$v_1(t) = A[1 + mf_1(t) + mf_2(t)] \cos \omega_0 t, \tag{10.1–20}$$

where A is the carrier amplitude and m is the modulation index. The vestigial sideband signal (which consists of f_1 normally modulated and f_2 single-sideband-modulated and increased by a factor of 2 relative to f_1) is given by

$$v_2(t) = A[1 + mf_1(t) + mf_2(t)] \cos \omega_0 t - Am\hat{f}_2(t) \sin \omega_0 t. \tag{10.1–21}$$

Since $|\hat{f}_2(t)|^2 \ll 1$, $v_2(t)$ is equivalent to an SSB signal with a large added carrier, and thus has an envelope of the form

$$A[1 + mf_1(t) + mf_2(t)] = A[1 + mf(t)], \tag{10.1–22}$$

which is identical with the envelope of the normal AM signal $v_1(t)$.

In practice, the amplitude of the upper-sideband information is not doubled at the transmitter; rather, the high-frequency portion of $f(t)$ is enhanced by a factor of 2 by the receiver after the signal has been demodulated. This not only simplifies the bandpass filter at the transmitter, which removes the lower-sideband information, but also permits faithful envelope demodulation of the vestigial sideband signal with the weaker requirement that $|\hat{f}_2(t)/2|^2 \ll 1$.

Peak Envelope Detection

An ideal peak envelope detector is a device which samples the peak of each positive (or negative) carrier cycle and holds the peak value until the next carrier cycle occurs. Figure 10.1–12 illustrates a typical set of input and output waveforms for an ideal peak envelope detector. It is apparent from Fig. 10.1–12 that a considerable amount of ripple appears on the output signal $v_o(t)$ unless the carrier frequency greatly exceeds the maximum frequency component ω_m of $g(t)$. Consequently, unless subsequent filtering is employed, the use of the envelope detector is restricted to situations where a very wide separation exists between ω_m and ω_0. However, when a wide separation between ω_m and ω_0 exists, it is apparent that $v_o(t)$ closely approaches $g(t)$ for $g(t) \geq 0$.

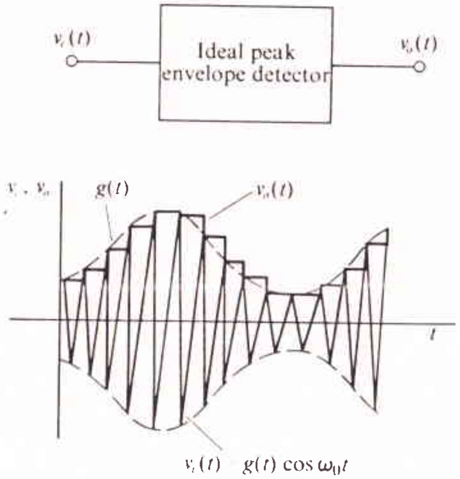

$v_i(t)$
Ideal peak envelope detector
$v_o(t)$
v_i, v_o
$g(t)$
$v_o(t)$
t
$v_i(t) \quad g(t) \cos \omega_0 t$

Fig. 10.1–12 Ideal peak envelope detector.

Most practical peak envelope detectors employ a diode to drive the holding network (usually a resistor in parallel with a capacitor) to the peak value of each carrier cycle, as shown in Fig. 10.1–13. Once $v_o(t)$ has reached the peak value of $v_i(t)$, the diode becomes reverse biased and $v_o(t)$ decays slowly toward zero with a time constant $\tau = RC$ until, near the peak of the cycle, $v_i = v_o$, which again turns the diode on and brings $v_o(t)$ to the peak value of $v_i(t)$. The resistor R in the holding network obviously has the effect of increasing the ripple; however, it is required in most practical detectors to ensure that $v_o(t)$ decays more rapidly during every holding period than the envelope of $v_i(t)$. If the decay in $v_o(t)$ is insufficient, the diode does not turn on at the peak of every cycle of $v_i(t)$, and "failure-to-follow" distortion results. Clearly the time constant τ must be chosen to meet a compromise between ripple and "failure-to-follow" distortion.

It is apparent that the peak envelope detector, like the average envelope detector, produces an output proportional to $|g(t)|$ which results in distortion if $g(t)$ is not always positive; hence suppressed carrier AM demodulation and SSB demodulation are

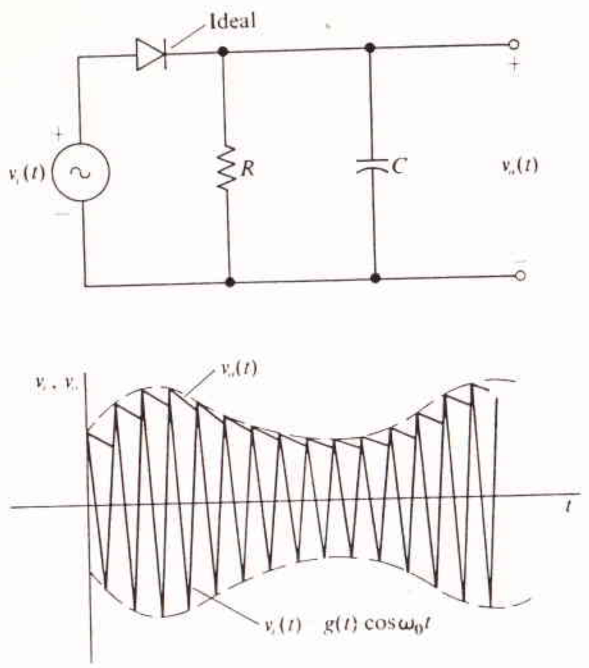


Fig. 10.1–13 Practical peak envelope detector.

impossible with the peak envelope detector. On the other hand, the peak envelope detector nicely demodulates normal AM signals, as well as vestigial sideband signals. Since exact expressions for $v_o(t)$ will be obtained in Sections 10.3 and 10.5 for specific practical peak detectors, no further general analysis is attempted at this point.

10.2 PRACTICAL AVERAGE ENVELOPE DETECTORS

In this section we consider some of the problems encountered in implementing circuits which achieve average envelope detection. Unfortunately, the ideal diode does not exist in nature; therefore, the greatest problem in designing an average envelope detector is the synthesis of the half-wave rectifier portion with physical diodes. We first show the effects on the detector of employing a physical diode in place of an ideal diode in the half-wave rectifier circuit; then we explore a method by which a physical diode can be made to function as an ideal diode in the detector circuit. This method utilizes feedback to remove the nonlinearities created by the nonideal diode from the half-wave rectifier characteristic.

To observe the basic problem which arises when physical diodes are employed in place of ideal diodes, we consider the detector shown in Fig. 10.2–1. If the diode D is ideal, then

$$i_D = \frac{v_i}{R} \qquad \text{for } v_i > 0$$

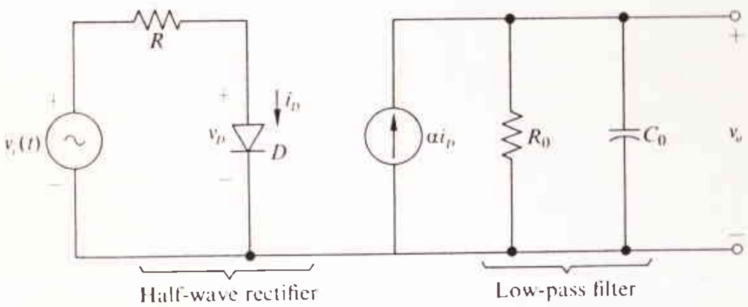


Fig. 10.2–1 Implementation of average envelope detector.

and

$$i_D = 0 \qquad \text{for } v_i \le 0;$$

hence i_D is an exact half-wave rectified version of v_i/R. This current, multiplied by α, is then placed through a low-pass filter (R_0 in parallel with C_0) to extract its average value. Consequently, if $v_i(t) = g(t)\cos\omega t$, then

$$v_o(t) = \frac{\alpha}{\pi R} g(t) * z_L(t). \tag{10.2–1}$$

where $z_L(t)$ is the response of $v_o(t)$ when $\alpha i_D(t) = \delta(t)$ (a unit impulse). If the filter is wide enough to pass $g(t)$, then, of course,

$$v_o(t) = \frac{\alpha R_0}{\pi R} g(t). \tag{10.2–2}$$

Equations (10.2–1) and (10.2–2) are just the results of the previous section applied to the circuit of Fig. 10.2–1.

However, if the diode is not ideal, the expressions for i_D and, in turn, v_o are modified. In particular, if the diode is modeled by the relationship

$$i_D = I_S(e^{qv_D/kT} - 1) \approx I_S e^{qv_D/kT}, \tag{10.2–3}$$

which is a reasonable approximation for most semiconductor diodes, then i_D may be expressed in terms of $v_i(t)$ by the equation

$$v_i = i_D R + \frac{kT}{q} \ln \frac{i_D}{I_S}. \tag{10.2–4}$$

Equation (10.2–4) may be also written in the normalized form

$$z = y + \ln y - \ln W, \tag{10.2–5}$$

where $z = qv_i/kT$, $y = qi_D R/kT$, and $W = qI_S R/kT$. A plot of y vs. z with $-\ln W$ as a parameter is shown in Fig. 10.2–2 along with the corresponding plot for an ideal diode. As $-\ln W$ increases, the entire z-y characteristic shifts linearly to the right. Here we see that the physical diode does not turn on at $v_i = 0$ (or $z = 0$), but rather requires a more positive value of v_i before any significant diode current flows. The effect of this turn-on requirement is to prevent small input signals from producing output. Thus such a detector cannot possibly demodulate 100% modulated als without distortion.

The turn-on requirement becomes more obvious if the actual half-wave rectifier racteristic is modeled by the two-segment piecewise-linear characteristic shown ig. 10.2–2 and given by

$$y = \begin{cases} z + \ln W - 5, & z > -\ln W + 5, \\ 0, & z \le -\ln W + 5, \end{cases} \tag{10.2–6}$$

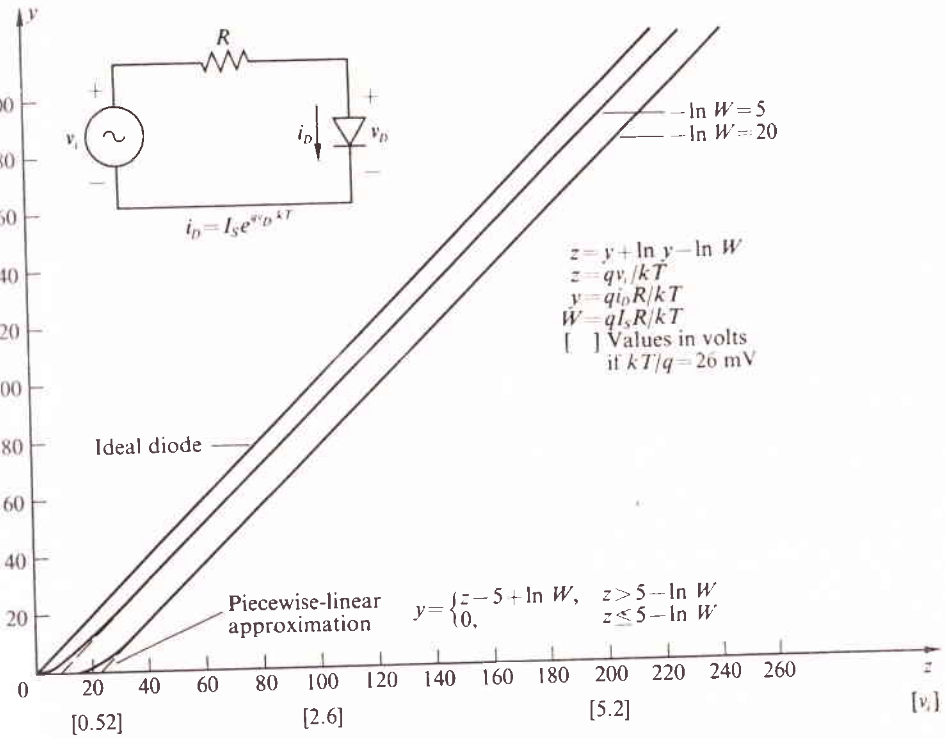


Fig. 10.2–2 Plot of $y\,[i_DR]$ vs. $z\,[v_i]$ with $-\ln W$ as a parameter.

r equivalently, by

$$i_D = \begin{cases} \dfrac{v_i - V_0}{R}, & v_i > V_0, \\ 0, & v_i \le V_0, \end{cases} \tag{10.2–7}$$

vhere

$$V_0 = \frac{kT}{q}\left(5 - \ln\frac{qI_SR}{kT}\right).$$

It should be apparent from Eq. (10.2–7) that constructing the piecewise-linear characteristic is equivalent to replacing the physical diode in the circuit shown in Fig. 10.2–1 by an ideal diode in series with a battery of value V_0, as shown in Fig. 10.2–3. From the equivalent circuit it is obvious that, unless $v_i(t)$ exceeds V_0, no output exists.

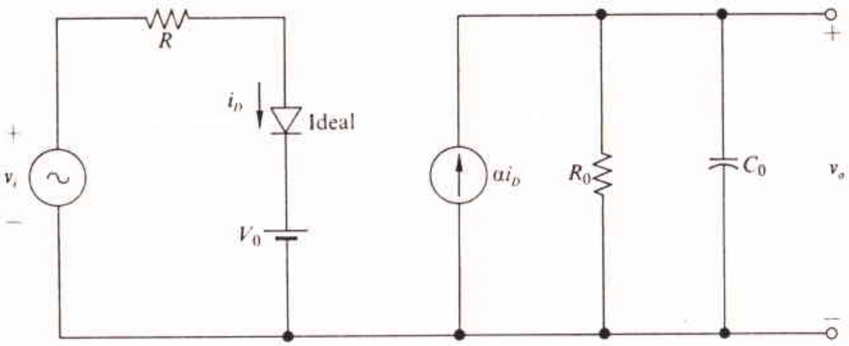


Fig. 10.2–3 Model for physical diode in average envelope detector circuit.

We should note that the turn-on voltage V_0 is a function of kT/q, R, and I_S. For example, with a germanium diode for which $I_S \approx 2 \times 10^{-7}$ A, $R = 1$ kΩ, and $kT/q = 26$ mV,

$$-\ln W = 4.86 \qquad \text{and} \qquad V_0 = 256\text{ mV}.$$

On the other hand, with a silicon diode for which $I_S \approx 2 \times 10^{-12}$ A, $R = 1$ kΩ, and $kT/q = 26$ mV,

$$-\ln(W) = 16.4 \qquad \text{and} \qquad V_0 = 556\text{ mV}.$$

It is apparent that the larger the values of R and I_S, the lower the value of V_0 and the smaller the departure of the physical half-wave rectifier characteristic. Clearly, then, a germanium diode with a large value of R provides the best approximation to the ideal half-wave rectifier.

To observe the effect of the turn-on voltage V_0 on the output of the circuit of Fig. 10.2–3, consider $v_i(t)$ to be equal to $V_1 \cos \omega t$. The diode current therefore has the form of a periodic train of sine-wave tips of peak value $I_{DP} = (V_1 - V_0)/R$ and conduction angle $2\phi = 2\cos^{-1}(V_0/V_1)$. Consequently the average value of i_D, which is extracted by the output filter, takes the form (cf. the appendix to Chapter 4)

$$\begin{aligned} \overline{i_D} &= \frac{I_{DP}}{\pi}\,\frac{\sin\phi - \phi\cos\phi}{1 - \cos\phi} \\ &= \frac{V_0}{\pi R}\left[\sqrt{\left(\frac{V_1}{V_0}\right)^2 - 1} - \cos^{-1}\frac{V_0}{V_1}\right] \\ &= \frac{V_0}{R}F\left(\frac{V_1}{V_0}\right), \end{aligned} \tag{10.2–8}$$

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA

ere

$$F\left(\frac{V_1}{V_0}\right) = \frac{1}{\pi}\left[\sqrt{\left(\frac{V_1}{V_0}\right)^2 - 1} - \cos^{-1}\frac{V_0}{V_1}\right].$$

lot of $F(V_1/V_0)$ appears in Fig. 10.2–4, from which it is observed that, for $V_1 > 4V_0$, V_1/V_0) may be approximated by its asymptotic value

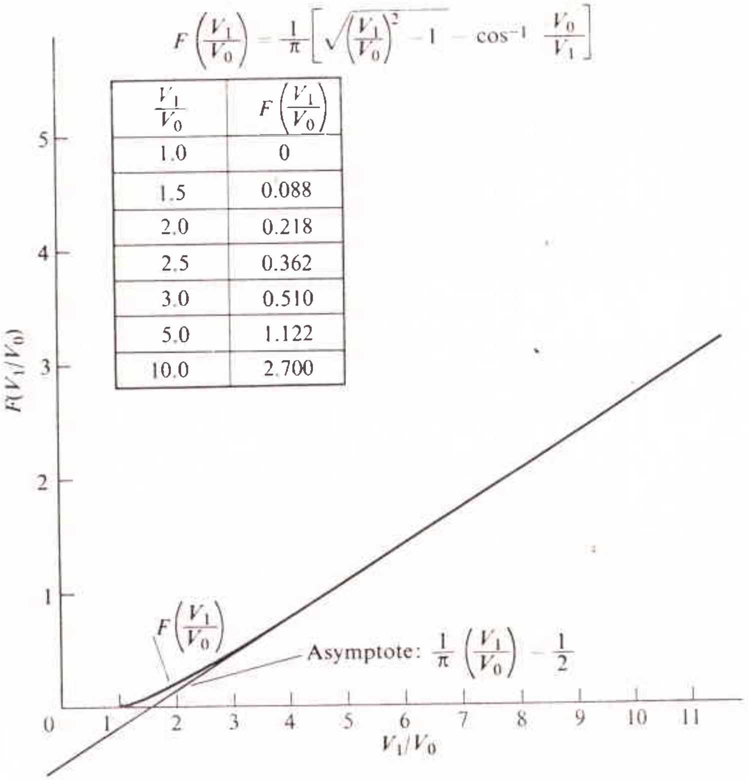
$$F\left(\frac{V_1}{V_0}\right) = \frac{1}{\pi}\left(\frac{V_1}{V_0}\right) - \frac{1}{2}. \tag{10.2–9}$$

us, for $V_1 > 4V_0$,

$$\overline{i_D} = \frac{V_1}{\pi R} - \frac{V_0}{2R} \tag{10.2–10}$$

d, if the output filter of the average envelope detector shown in Fig. 10.2–3 removes nal components at ω and its harmonics while being wide enough to pass the riations in V_1, then v_o is given by

$$v_0 = \alpha\overline{i_D}R_0 = \frac{\alpha V_1 R_0}{\pi R} - \frac{\alpha V_0 R_0}{2R}. \tag{10.2–11}$$



$\frac{V_1}{V_0}$	$F\left(\frac{V_1}{V_0}\right)$
1.0	0
1.5	0.088
2.0	0.218
2.5	0.362
3.0	0.510
5.0	1.122
10.0	2.700

Fig. 10.2–4 Plot of $F(V_1/V_0)$ vs. V_1/V_0.

If $V_1 = g(t)$, then

$$v_o(t) = \frac{\alpha g(t)R_0}{\pi R} - \frac{\alpha V_0 R_0}{2R}. \tag{10.2–12}$$

Thus the effect of the turn-on voltage is to subtract a constant value, $\alpha V_0 R_0/2R$ in this case, from the output, provided that $g(t) > 4V_0$. If $g(t)$ drops below $4V_0$, output distortion occurs, since $\overline{i_D}$ and $v_o(t)$ are nonlinearly related to $g(t)$, as can be seen from Fig. 10.2–4.

The restriction $g(t) > 4V_0$ for all t severely limits the maximum modulation index of the AM signal that may be linearly demodulated by the detector of Fig. 10.2–3. Consider, for example,

$$g(t) = V_1[1 + mf(t)]$$

where $|f(t)|_{\max} = 1$ and thus m is the modulation index. For this case, $g(t) > 4V_0$ is equivalent to the condition

$$V_1 > \frac{4V_0}{1 - m}. \tag{10.2–13}$$

Consequently, if $V_0 = 250$ mV and $m = 0.8$, V_1 must be greater than 5 V to achieve linear demodulation. As a practical matter, for voice demodulation, the distortion would be within tolerable limits with V_1 as low as 1.5 V.

A complete circuit which has as its model the circuit shown in Fig. 10.2–3 is illustrated in Fig. 10.2–5. In this circuit the voltage supply V_{CC} supplies the bias to keep the collector-base junction of the transistor reverse biased, while C is a coupling capacitor which isolates $v_i(t)$ from the bias supply. The emitter-base junction acts as the half-wave rectifier, while the diode D provides a return path for the capacitor current and thus prevents clamping; i.e., if the average value of $v_i(t)$ is zero, v_C must charge to V_{CC} to keep $\overline{i_D} = \overline{i_E}$, which is necessary to keep the average capacitor current equal to zero. With $v_C = V_{CC}$, with the emitter current described by the expression

$$i_E = I_{ES}e^{qv_{EB}/kT}, \tag{10.2–14}$$

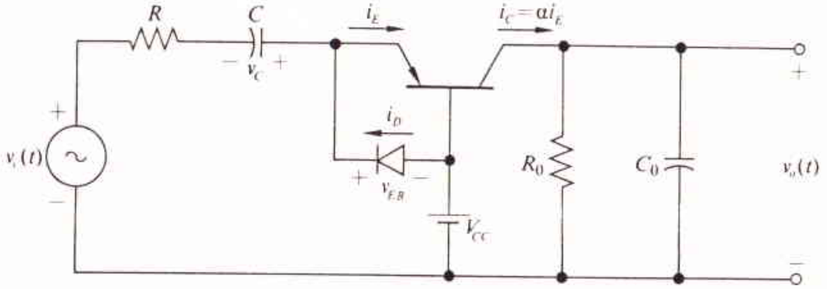


Fig. 10.2–5 Circuit for achieving average envelope detection.

nd with the assumption that C is a short to v_i, the emitter current may be related to $_i(t)$ by the expression

$$v_i(t) = Ri_E + \frac{kT}{q}\ln\frac{i_E}{I_{ES}}, \qquad (10.2\text{–}15)$$

vhich is the same relationship as that obtained from the circuit of Fig. 10.2–1; thus ve see that the emitter-base junction functions exactly as the diode in the previous roblem, while the collector current provides the current drive, αi_E, for the low-pass lter. Consequently, Eqs. (10.2–11) and (10.2–12) provide expressions for the output f the circuit of Fig. 10.2–4.

In choosing components for the circuit of Fig. 10.2–5 we usually choose two dentical germanium transistors and connect the base and collector of one together o obtain a diode. We then select a value of R somewhere in the vicinity of 10 kΩ and letermine V_0. A smaller value of R increases V_0, while a larger value of R requires $_i(t)$ to be too large to achieve a reasonable output current. For any given modulation ndex the required amplitude V_1 of

$$v_i(t) = V_1[1 + mf(t)]\cos\omega t$$

nay be determined to satisfy Eq. (10.2–13).

With the assumption that V_{CC} is also specified, R_0 is chosen to produce a reason-bly large output swing without causing the transistor to saturate. The capacitor $_0$ is chosen such that $1/R_0C_0 = \omega_3 \geq \omega_m$, where ω_m is the maximum modulation requency and ω_3 is the -3 dB bandwidth of the low-pass filter. Finally ωC is hosen to be at least ten times as great as $1/R$.

xample 10.2–1 For the average envelope detector shown in Fig. 10.2–6, determine values for all unspecified parameters such that linear demodulation is achieved. lso determine an expression for $v_o(t)$.

olution. Since $R = 10$ kΩ and $I_S = 2 \times 10^{-7}$ A, the turn-on voltage V_0 has the value $V_0 = 197$ mV; hence with a modulation index of 0.6 linear demodulation equires $V_1 > 1.97$ V (cf. Eq. 10.2–13). We shall choose $V_1 = 2.5$ V. With this choice.

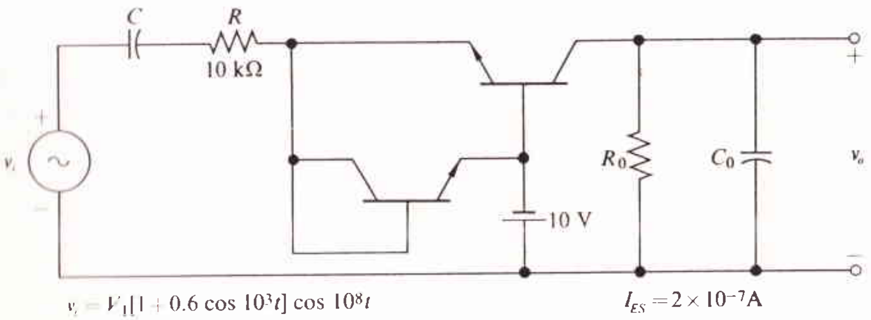

C
R
10 kΩ
+
v_i
−
R_0
C_0
10 V
+
v_o
−
$v_i = V_1[1 + 0.6\cos 10^3t]\cos 10^8t$
$I_{ES} = 2 \times 10^{-7}$A

Figure 10.2–6

$v_o(t)$ is given by (cf. Eq. 10.2–12)

$$v_o = -(250\,\mu\text{A})(1 + 0.6\cos 10^3t)\frac{R_0}{\pi} + (9.9\,\mu\text{A})R_0$$

with the assumptions that the transistor does not saturate, that $\alpha = 1$, and that the output filter passes frequencies up to 10^3 rad/sec undistorted. The minus sign, of course, results from employing an NPN transistor instead of a PNP transistor.

Since the minimum value of v_o must be greater than -10.2 V to avoid saturating the germanium transistor, we obtain

$$\left[(9.9\,\mu\text{A}) - \frac{(250\,\mu\text{A})1.6}{\pi}\right]R_0 > -10.2\text{ V},$$

or equivalently, $R_0 < 87$ kΩ. To leave some margin for parameter variation, we choose $R_0 = 68$ kΩ (a standard 5% value) and obtain finally

$$v_0 = -(4.7\text{ V}) - (3.2\text{ V})\cos 10^3t.$$

By choosing $1/R_0C_0 = 10^4$ rad/sec or equivalently $C_0 \approx 1500$ pF, we ensure that the signal information is transmitted and that frequency components at the fundamental and harmonics of the carrier frequency are removed from the output. At $\omega = 10^8$ rad/sec, $1/10^8C_0 = 6.8\ \Omega$, which is indeed a short circuit compared with 68 kΩ. If the impedance of the output filter at the carrier frequency is not small compared with R_0, ripple appears in the output. Although this ripple may be removed by subsequent filtering, care must be taken to ensure that it does not saturate the transistor when added to v_o. Finally, to ensure that the coupling capacitor is an ac short circuit, we need $C \geq 10/\omega R = 10$ pF. Again, to allow for a margin of safety, we choose $C = 100$ pF.

The nonideal effect of the physical diode may be greatly reduced by constructing a half-wave rectifier that incorporates the diode in the feedback loop of an operational

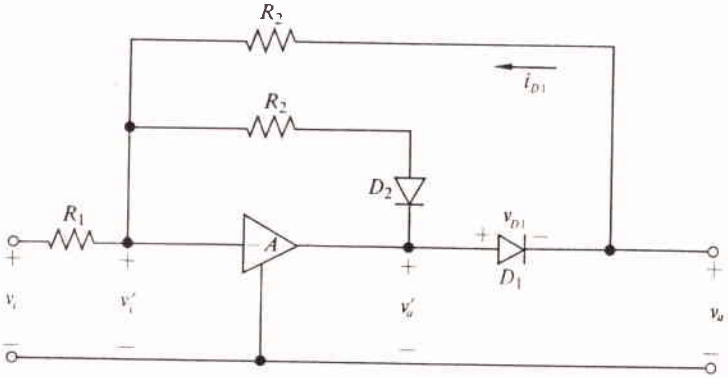


Fig. 10.2–7 Half-wave rectifier employing a diode in the feedback loop of an operational amplifier.

amplifier, as shown in Fig. 10.2–7. The half-wave rectification is achieved with diode D_1, which conducts for negative values of v_i and opens for positive values of v_i. Diode D_2 is incorporated in the circuit to prevent open-loop operation and the possibility of device saturation for positive values of v_i for which D_1 opens. Since with D_1 and D_2 in the circuit the maximum value of v_a' exceeds the maximum value of v_a by only the small diode "on" voltage (approximately 220 mV for germanium and 650 mV for silicon—cf. Section 5.4), the maximum value of v_i' is closely approximated by $v_{a_{\max}}/A$, where A is the voltage amplification of the operational amplifier. In addition, if the positive and negative peak values of v_i are of the same order of magnitude, the presence of D_2 ensures that the minimum value of v_i' is of the order of $-v_{a_{\max}}/A$. Consequently, for $A > 100$ (which is true even for the poorest-quality operation amplifier), v_i' appears as a "virtual" ground relative to v_a, and thus v_a may be very closely approximated by

$$v_a = i_{D1}R_2. \tag{10.2–16}$$

Noting that D_2 is open for $i_{D1} > 0$, we may relate i_{D1} to v_1 by the expression

$$-v_i = i_{D1}\left(R_1 + \frac{R_2}{1 + A}\right) + \frac{kT}{q(1 + A)} \ln \frac{i_{D1}}{I_S}. \tag{10.2–17}$$

In deriving Eq. (10.2–17), we assume that the diode current and voltage are again related by the approximation of Eq. (10.2–3), that the operational amplifier input impedance is several orders of magnitude greater than R_1 or R_2 and thus may be neglected, and that the operational amplifier output impedance is small in comparison with R_2. These assumptions are usually true for most integrated operational amplifiers; and even in those cases where they are not strictly true, it can be readily shown that the results we derive are still valid. The equations, however, become more cumbersome.

Equation (10.2–17) may be written in the normalized form

$$z = y + \frac{1}{1 + A}(\ln y - \ln W), \tag{10.2–18}$$

where

$$z = -\frac{qv_i}{kT}, \qquad y = \frac{qi_{D1}}{kT}\left(R_1 + \frac{R_2}{1 + A}\right), \qquad \text{and} \qquad W = \frac{qI_S}{kT}\left(R_1 + \frac{R_2}{1 + A}\right).$$

It is apparent that, for $A = 0$, Eq. (10.2–18) reduces to Eq. (10.2–5) and no improvement over the previous half-wave rectifier is obtained. As A is increased from zero, the effect of the nonlinear term in y is reduced and the z-y characteristic approaches that of an ideal diode. A plot of y vs. z for $-\ln W = 20$ is shown in Fig. 10.2–8. The scale in this figure is greatly expanded from the scale of Fig. 10.2–2, as can be seen from the z-y characteristic for $A = 0$. From Fig. 10.2–7 we observe that for $A \geq 99$ the z-y characteristic is within 7 mV (unnormalized) of the ideal half-wave rectifier characteristic ($A \to \infty$). In addition, for $A \geq 999$ the characteristic is within 0.7 mV of the ideal half-wave rectifier characteristic and for all practical purposes may be considered to be ideal, that is,

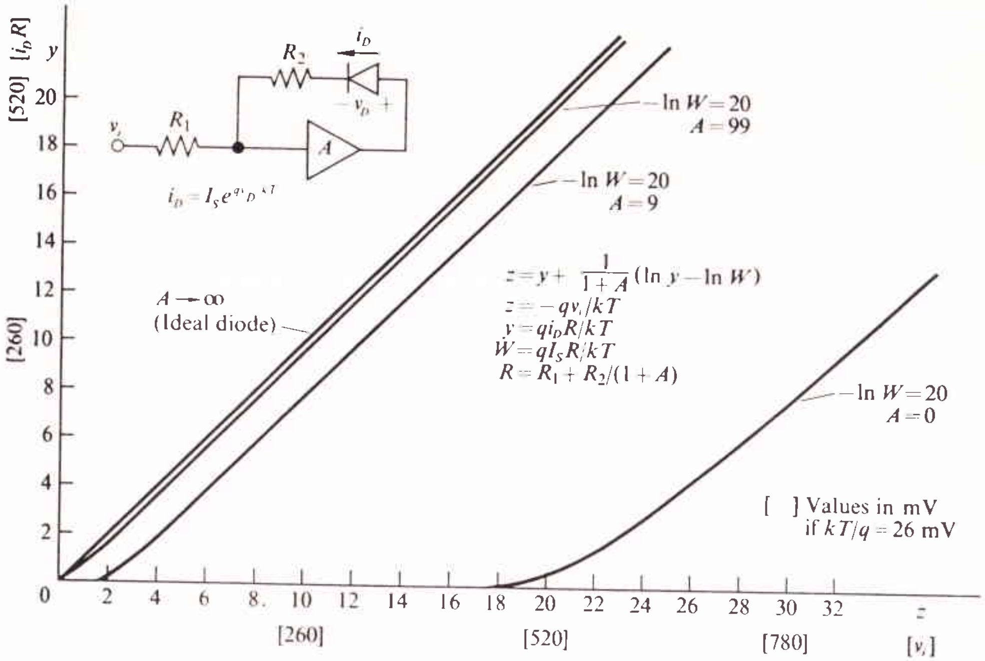


Fig. 10.2–8 Plot of y vs. z with $-\ln W = 20$ and with A as a parameter.

$$y = \begin{cases} z, & z > 0, \\ 0, & z \leq 0, \end{cases} \tag{10.2–19}$$

or equivalently,

$$i_{D1} = \begin{cases} \dfrac{-v_i}{R_1 + R_2/(1 + A)} \approx \dfrac{-v_i}{R_1}, & v_i < 0, \\ 0, & v_i \geq 0. \end{cases} \tag{10.2–20}$$

Clearly, then, v_a is related to v_i by the ideal half-wave rectifier characteristic given by

$$v_a = \begin{cases} \dfrac{-v_i R_2}{R_1}, & v_i < 0, \\ 0, & v_i \geq 0. \end{cases} \tag{10.2–21}$$

If hot-carrier diodes are used for D_1 and D_2, Eq. (10.2–21) provides an accurate description of the transfer characteristic of v_a vs. v_i for carrier frequencies up to the tens of megahertz, provided that at these frequencies the open-loop amplification of the operational amplifier still exceeds 100.

A complete average envelope detector employing two operational amplifiers is shown in Fig. 10.2–9. If $R_3 \gg R_2$, which should be the case to ensure that the

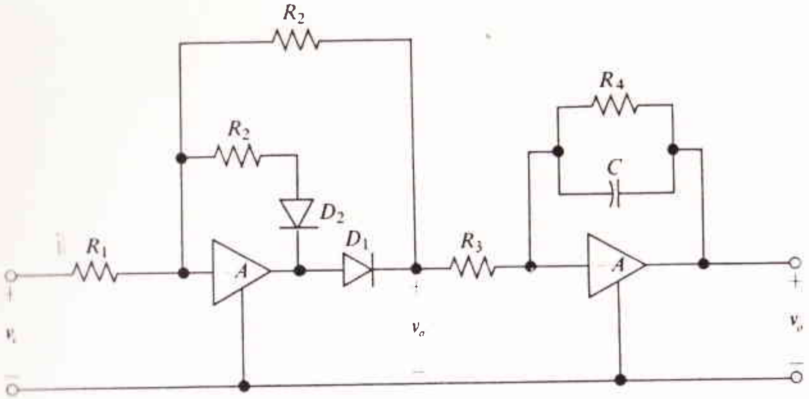


Fig. 10.2–9 Average envelope detector employing operational amplifiers.

current from D_1 flows primarily through R_2, then the first amplifier performs as an ideal half-wave rectifier, while the second amplifier performs as a low-pass filter with a transfer function

$$H(p) = \frac{V_0(p)}{V_a(p)} = \frac{R_4/R_3}{1 + pR_4C}. \tag{10.2–22}$$

Equation (10.2–22) is the transfer function of a low-pass filter with a transmission at low frequencies of R_4/R_3 and a -3 dB bandwidth of $\omega_3 = 1/R_4C$. Therefore, if $v_i(t) = g(t)\cos\omega t$, then

$$v_o(t) = \frac{R_2}{\pi R_1} g(t) * h(t), \tag{10.2–23}$$

where $h(t) = \mathscr{L}^{-1}H(p)$. If ω_3 is large in comparison with the maximum frequency component of $g(t)$, then

$$v_o(t) = \frac{R_2R_4}{\pi R_1R_3} g(t). \tag{10.2–24}$$

By correctly adjusting R_1, R_2, R_3, and R_4, any desired amplification may be achieved provided that the operational amplifiers are not driven into saturation. In addition, by cascading several more operational amplifiers, a higher-order, low-pass filter can readily be synthesized.

10.3 NARROWBAND PEAK ENVELOPE DETECTOR

One of the most widely used demodulators for normal AM signals is the narrowband peak envelope detector shown in Fig. 10.3–1. Almost every superheterodyne AM receiver has a detector of this form, which is constructed by placing the series combination of a diode D and a parallel R_0-C_0 circuit across the output-tuned circuit of the last IF amplifier in the IF strip. The parallel RLC circuit driven by the current source $i_i(t)$ provides the model for the final IF amplifier. The narrowband peak envelope detector is always designed so that the loaded Q_T of the parallel RLC circuit is high and the time constant of the parallel R_0-C_0 circuit is long in comparison with the period of a carrier cycle. The high value of Q_T ensures that the last IF stage is sufficiently selective and also prevents the nonlinear diode circuit from distorting the waveform of $v_i(t)$. The long R_0C_0 time constant enables $v_o(t)$ to remain essentially constant at the peak value of $v_i(t)$ over each carrier cycle. Thus output ripple is minimized.

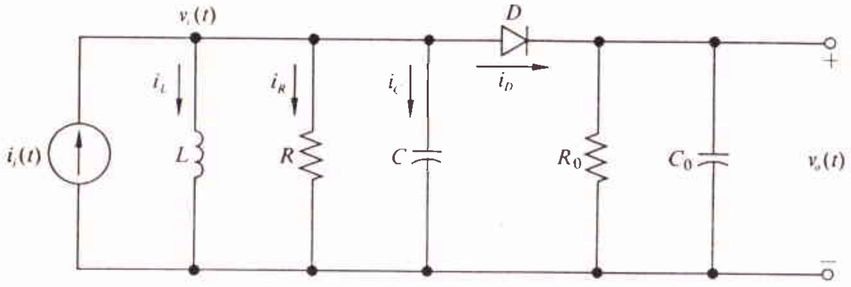


Fig. 10.3–1 Narrowband peak envelope detector.

In analyzing the detector of Fig. 10.3–1 we assume that the diode is ideal. This assumption greatly simplifies the analysis and, in addition, as we shall demonstrate, provides results that are directly applicable to the identical circuit employing a physical diode.

We begin the analysis by obtaining an expression for $v_i(t)$ and $v_o(t)$ for the case where

$$i_i(t) = I_1 \cos\omega_0 t.$$

In particular, we show that $v_o(t)$ is indeed directly proportional to I_1, as is required by a linear envelope detector. We then generalize the analysis to the dynamic case, where

$$i_i(t) = b(t)\cos\omega_0 t = I_1[1 + mf(t)]\cos\omega_0 t,$$

and show that a simple equivalent circuit exists for determining $v_o(t)$ as a function of $b(t)$. Finally, we consider the problem of "failure-to-follow" distortion which results when the detector is improperly designed.

Static Analysis

If the R_0C_0 time constant is long in comparison with the duration $T = 2\pi/\omega_0$ of a carrier cycle (or, equivalently, C_0 is an ac short circuit at ω_0), then with $i_i(t) = I_1\cos\omega_0 t$, the output voltage $v_o(t)$ stabilizes at a dc value $v_o = V_{dc}$. With V_{dc} developed across the capacitor C_0, it is apparent from Example 5.5–3 that, if the loaded Q_T of the parallel RLC circuit is high, then the voltage across the tuned circuit $v_i(t)$ must have the form

$$v_i(t) = V_{dc}\cos\omega_0 t. \tag{10.3–1}$$

To determine V_{dc} we observe that the diode D is reverse biased except in the nmediate vicinity of the peak of $v_i(t)$; hence the diode current i_D must flow in narrow ulses occurring at the peak of each cycle of $v_i(t)$, as shown in Fig. 10.3–2. Conseuently, i_D may be expanded in a Fourier series of the form

$$i_D(t) = I_{D0}(1 + 2\cos\omega_0 t + \cdots), \tag{10.3–2}$$

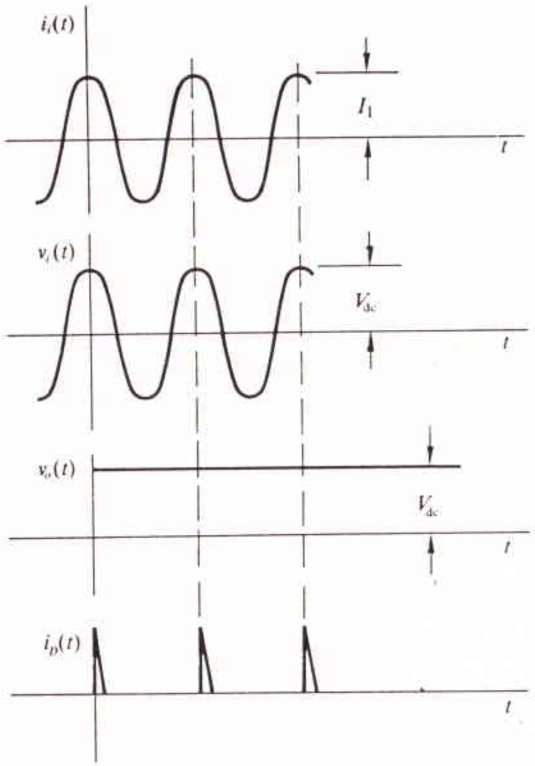


Fig. 10.3–2 Current and voltage waveforms occurring in the circuits of Fig. 10.3–1.

where I_{D0} is the average value of $i_D(t)$. [The narrow pulse width causes the amplitude of the fundamental component of $i_D(t)$ to be twice the average value.] Since the average diode current must flow through R_0,

$$V_{dc} = I_{D0}R_0. \tag{10.3–3}$$

In addition, by applying Kirchhoff's current law to the fundamental current components flowing into the parallel RLC circuit, we obtain

$$I_1 = \frac{V_{dc}}{R} - 2I_{D0}; \tag{10.3–4}$$

and by combining Eqs. (10.3–3) and (10.3–4), we finally obtain the desired expression for V_{dc}:

$$V_{dc} = I_1 R_T, \tag{10.3–5}$$

where $R_T = R \| (R_0/2)$. We also observe that the equivalent linear loading on the tuned circuit produced by the nonlinear diode circuit is

$$\frac{V_{dc}}{2I_{D0}} = \frac{R_0}{2} \tag{10.3–6}$$

and that Q_T is given by

$$Q_T = \frac{R_T}{\omega_0 L} = R_T \omega_0 C. \tag{10.3–7}$$

It is quite apparent from Eq. (10.3–5) that V_{dc} is indeed linearly related to I_1 and, therefore, that the circuit of Fig. 10.3–1 functions, at least statically, as an envelope detector. To understand the operation of the circuit still better, we determine the form of $i_D(t)$. If we assume that the diode becomes forward biased at t_1 and becomes reverse biased at t_2, where the time interval $t_0 = t_2 - t_1$ is short in comparison with T and occurs in the vicinity of the peak of $v_i(t)$, and if in addition we assume that $v_o(t)$ remains essentially constant at V_{dc} during diode conduction, then for $t_1 < t \leq t_2$

$$i_i \approx I_1,^{\dagger}$$

$$i_R = \frac{V_{dc}}{R} = i_{R_0},$$

$$i_L = i_L(t_1) + \frac{V_{dc}(t - t_1)}{L}. \tag{10.3–8}$$

Since all the nonreactive shunt branches of the circuit of Fig. 10.3–1 contain constant currents for $t_1 < t \leq t_2$, the ac ramp component of i_L must divide between C and C_0; however, the ramp component through C_0 also flows through the diode. Thus the diode current consists of a ramp with slope $-V_{dc}C_0/L(C_0 + C)$, which reaches 0 for $t = t_2$ (the diode becomes reverse biased when its current reaches zero). A sketch of $i_D(t)$ is shown in Fig. 10.3–3, which indicates that $i_D(t)$ consists of a periodic train of narrow triangular pulses.

The pulse width t_0 is readily determined by noting from Fig. 10.3–3 that the average value, I_{D0}, of $i_D(t)$ may be expressed as

$$I_{D0} = \frac{V_{dc}t_0^2 C_0}{2LT(C_0 + C)}, \tag{10.3–9}$$

---

† A sine wave remains essentially constant in the vicinity of its peak.

while from Eq. (10.3–3) $I_{D0} = V_{dc}/R_0$; hence eliminating I_{D0} from Eq. (10.3–12) yields

$$\frac{t_0}{T} = \sqrt{\frac{2L(C_0 + C)}{TR_0C_0}} = \sqrt{\frac{\omega_0 L(C_0 + C)}{\pi R_0 C_0}}$$

$$= \sqrt{\frac{C_0 + C}{\pi R_0 \omega_0 C_0 C}} = \sqrt{\frac{1}{\pi Q_{TS}}}, \qquad (10.3\text{–}10)$$

where $Q_{TS} = \omega_0 C_s R_0$ and $C_s = CC_0/(C + C_0)$ is the series combination of C and C_0.

In almost all narrowband detectors of interest, Q_{TS} is slightly greater than Q_T; hence a high value of Q_T ensures a high value of Q_{TS} and, in turn, a narrow pulse width. For example, if $C_0 = C$, then

$$Q_{TS} = \left(1 + \frac{R_0}{2R}\right)Q_T; \qquad (10.3\text{–}11)$$

if $C_0 \gg C$, then

$$Q_{TS} = \left(2 + \frac{R_0}{R}\right)Q_T. \qquad (10.3\text{–}12)$$

If $Q_{TS} = 200$, then $t_o/T = \phi/2\pi = 0.0399$ and $\phi = 14.4°$, where ϕ is the conduction angle. If $Q_{TS} = 100$, $\phi = 20.3°$. For Q_T as low as 50 the conduction angle is sufficiently small to justify the assumptions that the amplitude of the fundamental component of i_D is twice the average value and that i_i may be approximated by its peak value during the interval t_0. In addition, with a conduction angle less than 30°, the small flat occurring on the peak of $v_i(t)$ due to the forward bias of the diode is hardly perceptible.

If the R_0C_0 time constant is not infinite, v_0 increases slightly during the time that the diode is forward biased and decreases during the time the diode is reverse biased, thus causing a small amount of ripple on the output. It is readily shown that, even with a small amount of ripple, Fig. 10.3–3 and Eq. (10.3–10) still provide very good approximations for the diode current waveform and the current pulse duration. In particular, if v_0 increases during diode conduction, then the current pulse duration is slightly shorter than the duration given by Eq. (10.3–10); this gives even more validity to the previous assumptions.

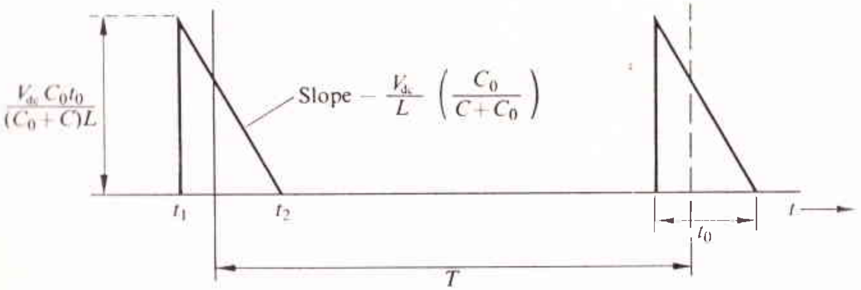


Fig. 10.3–3 Sketch of $i_D(t)$ vs. t.

To get some idea of the size of the increase ΔV in v_o during diode conduction, we observe that the net charge entering C_0 during a cycle must equal zero. With the assumption that $t_0 \ll T$, the charge entering the capacitor is $Q_+ = C_0\Delta V$, while the charge leaving is $Q_- \approx V_{dc}T/R_0$. (The approximation becomes exact as $t_0 \to 0$.) Consequently, the fractional ripple $\Delta V/V_{dc}$ appearing at the envelope detector output is closely approximated by

$$\frac{\Delta V}{V_{dc}} = \frac{T}{R_0C_0} = \frac{T}{\tau_0}, \qquad (10.3\text{–}13)$$

where $\tau_0 = R_0C_0$. For 1% ripple, $\tau = 100T$. If the fractional ripple remains below 0.1, all the previous results obtained provide excellent approximations (within 5%) to the actual results; however, with $\Delta V/V_{dc} = 0.1$, some additional low-pass filtering must follow the detector to remove the excess ripple.

Example 10.3–1 For the circuit shown in Fig. 10.3–4, (a) determine the value of v_o and (b) determine the peak-to-peak ripple on v_o.

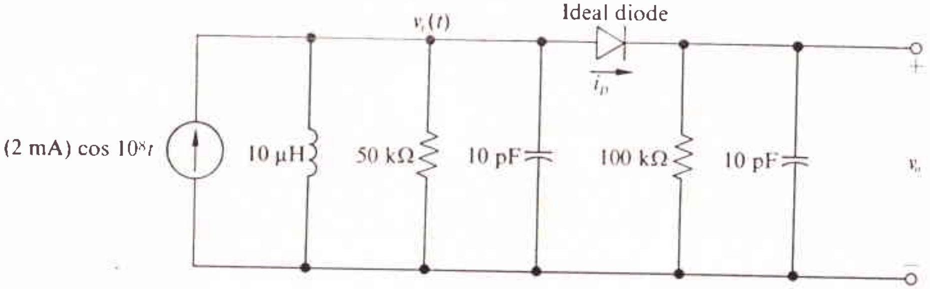


Figure 10.3–4

Solution. If we assume $Q_T > 10$, we can replace the diode and the circuitry to its right by an equivalent 50 kΩ resistor; hence

$$v_i(t) = (2\text{ mA})(25\text{ k}\Omega)\cos 10^8t = (50\text{ V})\cos 10^8t$$

and, in turn,

$$v_o(t) = (50\text{ V}) + \text{ripple}.$$

From Eq. (10.3–14) it is apparent that the fractional ripple has the value

$$\frac{\Delta V}{V_{dc}} = \frac{2\pi}{10^8}\,\frac{1}{(10\text{ pF})\times(100\text{ k}\Omega)} = 0.0628;$$

or equivalently, the peak-to-peak ripple has the value $\Delta V = 3.14$ V. It is also apparent that $Q_T = (25\text{ k}\Omega)/\omega_0 L = 25$, which justifies the initial assumption.

ynamic Analysis

'e now let

$$i_i(t) = I_1[1 + mf(t)]\cos\omega_0(t) = b(t)\cos\omega_0 t, \tag{10.3–14}$$

here $b(t) = I_1[1 + mf(t)]$, and assume again that the loaded Q_T is sufficiently high that $v_i(t)$ has frequency components only in the vicinity of ω_0 and may therefore written in the general form

$$v_i(t) = g(t)\cos\omega_0 t, \tag{10.3–15}$$

here $g(t)$ is some low-frequency waveform. If we assume, in addition, that the ripple n $v_o(t)$ is small and that the diode D conducts every carrier cycle, then $v_o(t)$ is iven by

$$v_o(t) = g(t). \tag{10.3–16}$$

Note that $v_o(t)$ cannot be less than $g(t)$, since this would require a voltage across a rward-biased ideal diode at the peak of $v_i(t)$ or a large amount of ripple. On the ther hand, $v_o(t)$ cannot be greater than $g(t)$ or the diode could not conduct every ycle; hence, with the above assumptions, $v_o(t) = g(t)$.] If the diode does not conduct very cycle, the result is failure-to-follow distortion, which we investigate later in his section.

Clearly, our objective is to evaluate $g(t)$ in terms of $b(t)$. To achieve this objective e first observe that a high value of Q_T ensures an even higher value of Q_{TS} and thus nsures that the diode current $i_D(t)$ flows in narrow pulses. Therefore, in this case, s in the static analysis, i_D may be expanded in a series of the form

$$i_D(t) = i_{D0}(t)(1 + 2\cos\omega_0 t + \cdots), \tag{10.3–17}$$

here $i_{D0}(t)$ is the slowly varying average value of $i_D(t)$. [The reader should convince imself that the expansion of Eq. (10.3–17) is a unique representation for $i_D(t)$ provided hat $i_D(t)$ is band-limited to $\omega_0/2$.] The slowly varying diode current flowing through he parallel R_0-C_0 circuit gives rise to $g(t)$; hence

$$g(t) = i_{D0}(t) * z_0(t), \tag{10.3–18}$$

where $z_0(t)$ is the response of the parallel R_0-C_0 circuit when driven by a unit impulse f current.

In addition, the input current $i_i(t)$, less the fundamental component of the diode urrent, flowing through the parallel RLC circuit gives rise to $v_i(t)$; that is,

$$v_i(t) = g(t)\cos\omega_0 t = [b(t)\cos\omega_0 t - 2i_{D0}(t)\cos\omega_0 t] * z(t), \tag{10.3–19}$$

where $z(t)$ is the impulse response of the parallel RLC circuit. By noting that $z(t)$ is symmetric narrowband filter with $\theta(\omega_0) = 0$, and by employing Eq. (3.3–3), we can simplify Eq. (10.3–19) to the form

$$g(t)\cos\omega_0 t = \{[b(t) - 2i_{D0}(t)] * z_L(t)\}\cos\omega_0 t, \tag{10.3–20}$$

or equivalently, to the form

$$g(t) = \left[\frac{b(t)}{2} - i_{D0}(t)\right] * 2z_L(t), \tag{10.3–21}$$

where $z_L(t)$ is the low-pass equivalent of $z(t)$. The response $z_L(t)$ is the impulse response of the circuit shown in Fig. 3.1–3. The response $2z_L(t)$ is the impulse response of a circuit with double the impedance level of the circuit shown in Fig. 3.1–3.

Equations (10.3–18) and (10.3–21) comprise two simultaneous equations from which $g(t)$ may be obtained. The method for obtaining $g(t)$ from these equations becomes apparent only when the two circuits described by these equations are drawn as in Fig. 10.3–5. Connecting terminals a–a' to terminals b–b' yields an equivalent circuit from which $g(t)$ may be obtained directly in terms of $b(t)$.

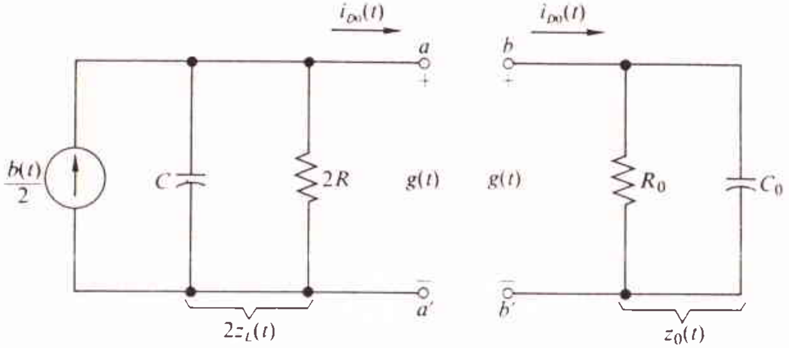


Fig. 10.3–5 Two circuits from which Eq. (10.3–18) (right) and Eq. (10.3–21) (left) can be obtained.

We note from the equivalent circuit of Fig. 10.3–5 that $g(t)$ may be related to $b(t)$ by the transfer function

$$Z_T(p) = \frac{G(p)}{B(p)} = \frac{R_T}{1 + p/\omega_3}, \tag{10.3–22}$$

where

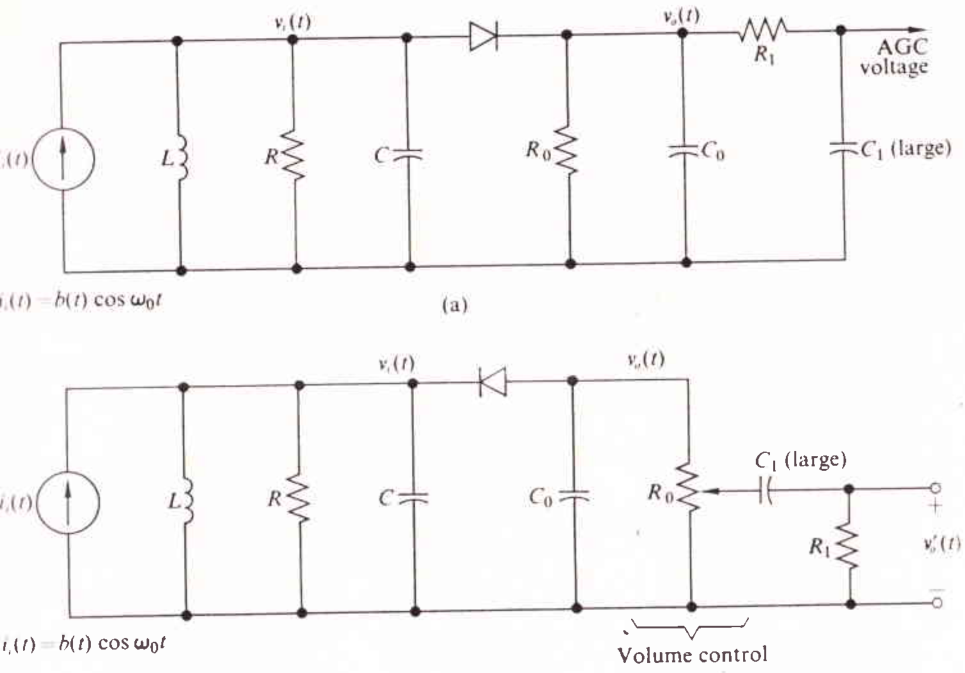
$$\omega_3 = \frac{1}{2R_T(C_0 + C)}, \qquad R_T = R \| \frac{R_0}{2},$$

and $G(p)$ and $B(p)$ are the Laplace transforms of $g(t)$ and $b(t)$ respectively. Here we see that the -3 dB bandwidth, which must be chosen large enough to pass $b(t)$, is a function of not only the parallel R_0-C_0 circuit but also the parallel RLC circuit. Hence in designing a narrowband peak envelope detector one must consider the complete equivalent circuit of Fig. 10.3–5. However, if ω_3 is wide enough to transmit the information of $b(t)$ undistorted, then

$$g(t) = b(t)R_T = I_1[1 + mf(t)]R_T \tag{10.3–23}$$

nd the desired envelope information is obtained. Clearly, Eq. (10.3–23) could have een obtained directly from Eq. (10.3–5) with the assumption of slowly varying modu- ation. For $C_0 \ll C$, Eq. (10.3–22) can be obtained by reflecting $R_0/2$ across the arallel RLC circuit (as was done in the static case) and then determining the relation- hip between g(t) and $b(t)$ for the resultant parallel RLC circuit.

If additional circuitry is added to the peak detector output for the purpose of leveloping an AGC voltage (Fig. 10.3–6a) or for the purpose of ac-coupling the de- nodulated signal to a subsequent stage (Fig. 10.3–6b), the previous analysis is still



(b)

Fig. 10.3–6 Additional circuits placed on detector output to (a) develop an AGC voltage and to (b) ac-couple the detector output.

valid provided that Q_T is high and the output ripple is small; i.e., the bandpass circuit to the left of the diode may be replaced by the equivalent circuit shown in Fig. 10.3–5 (left). A complete equivalent circuit for the envelope detector containing a coupling network (Fig. 10.3–6b) is shown in Fig. 10.3–7. In this figure the polarity of the current source has been reversed, since the diode in the original circuit is positioned to obtain the negative envelope. Two examples now demonstrate the usefulness of the equivalent circuits.

Example 10.3–2 For the narrowband peak envelope detector shown in Fig. 10.3–8, determine expressions for $v_o(t)$ and $v_i(t)$. Repeat the problem for $C_0 = 10\,\mu$F.

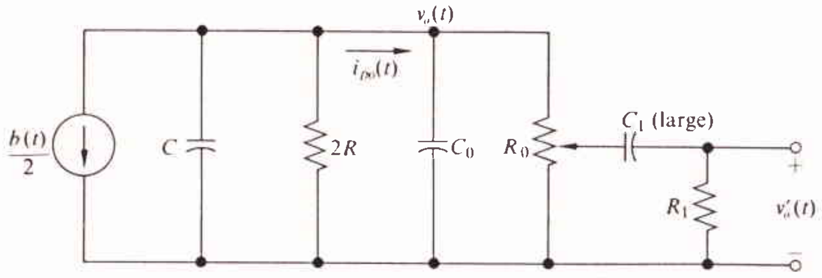


Fig. 10.3–7 Equivalent circuit for the detector shown in Fig. 10.3–6 (b).

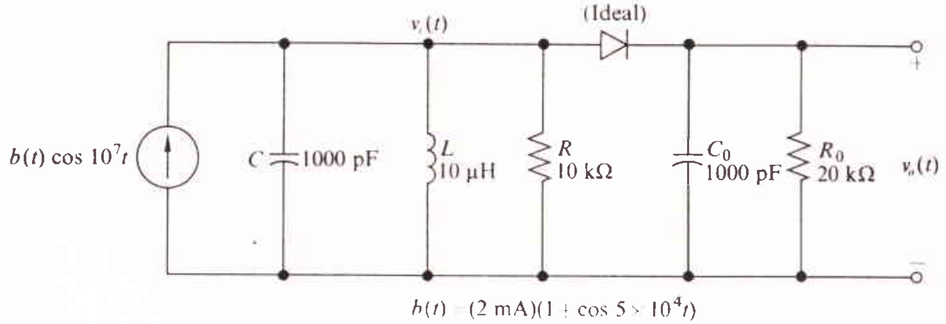


Figure 10.3–8

Solution. For the circuit of Fig. 10.3–8,

$$\omega_0 = \frac{1}{\sqrt{LC}} = 10^7 \text{ rad/sec} \qquad \text{and} \qquad Q_T = \frac{(R_0/2)\|R}{\omega_0 L} = 50.$$

In addition, the fractional ripple has the value

$$\Delta V/V_{\text{dc}} = 2\pi/\omega_0 R_0 C_0 = 0.0314;$$

hence, with a high value of Q_T, a small amount of ripple, and the assumption of no "failure-to-follow" distortion, the detector of Fig. 10.3–8 may be replaced by the equivalent circuit shown in Fig. 10.3–9. From the equivalent circuit it is apparent that

$$v_o(t) = (10\text{ V})\left[1 + \frac{1}{\sqrt{2}}\cos\left(5 \times 10^4 t - \frac{\pi}{4}\right)\right].$$

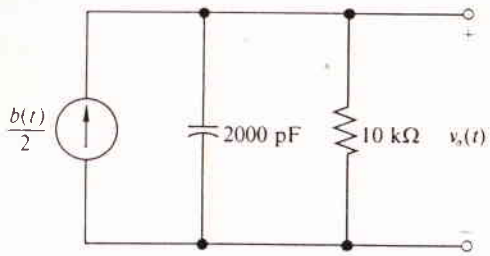
In addition, since $v_o(t)$ is equal to the envelope of $v_i(t)$,

$$v_i(t) = (10\text{ V})\left[1 + \frac{1}{\sqrt{2}}\cos\left(5 \times 10^4 t - \frac{\pi}{4}\right)\right]\cos 10^7 t.$$

We observe that, even though the input current has a modulation index of unity, the voltage developed across the tuned circuit has a modulation index of $1/\sqrt{2}$.

INDEX

Abramowitz, M., 641
AFC, 590
AGC, 308, 331–336, 486, 530
Alignability, 324–327
AM generators: balanced, 382–384
  differential amplifier, 374–375
  diode bridge, 376–379, 382–384
  FET, 380–381
  integrated circuit, 395–396
  square law, 384–387
  tuned circuit, 387–392
  See also Diode bridge, FET, Chopper modulation, Multiplier, *and/or* Modulation
AM methods: analog, 354–355
  chopper, 355–358
  direct, 361–362
  high level, 447–450
  nonlinear, 359–360
AM on FM signals, 589–590
AM receiver, superheterodyne, 478
AM signals: in capacitively coupled circuits, 168–169, 195–199
  in narrowband filters, 72, 76, 88–89
  normal, 347
  single sideband, 350–352, 457–461, 464–466
  suppressed carrier, 349, 357
Amplification, AM, 405–406
Amplifiers: cascode, 324–326
  class A, 401–405
  class AB, 401
  class B, broadband, 401, 421–425
  class B "linear," 92, 405–408
  class C "linear," 408–410
  class C–RF, 401, 410–415
  class D, broadband, 426–432
  class D, narrowband, 401, 415–421
  differential pair, 114–119, 324–326
  feedback, 206, 475–478
  FET, 98–104, 131–136
  gain controlled, 214, 335–336
  maximum signal without distortion, 4, 100, 110
  narrowband-tuned, 6–8, 314–331, 405–421
  power, 401–455
  small signal range, 4–5, 110, 118
  transistor, 1–8, 104–113
  tuned, 6–8, 314–331, 405–421
Amplitude control, 212–215, 265–266, 308, 331–336
Amplitude limiting, 7–8, 110, 188, 571–577
  nonlinear, 207
Amplitude modulation; *see specific headings under* AM
Amplitude sensitivity, function, 230–231, 264–265
Amplitude stabilization: FM, 530–532
  oscillators, 212–215, 262–266
Analog computer, 528
Angelo, E. J., Jr., 13
Armstrong, E. H., 560
Astable multivibrator, 553–556
Attenuator, nonlinear, 213, 265–266
Auto transformer, 49–51
Automatic frequency control (AFC), 590
Automatic gain control; *see* AGC
AVC; *see* AGC
Average envelope detection, 461–466
  in FM detectors, 582–584, 587
Average envelope detector, 468–478
  dual transistor, 473–475
  operational amplifier, 475–478
Avins, J., 613



$\frac{b(t)}{2} = (1\text{ mA})(1 + \cos 5 \times 10^4 t)$

Figure 10.3–9

If C_0 is replaced by a 10 μF capacitor, the capacitor in the equivalent circuit of ig. 10.3–9 must be replaced by a 10 μF + 1000 pF ≈ 10 μF capacitor. Such a arge capacitor effectively shorts the signal component of $b(t)/2$ to ground to yield $_o(t) = 10$ V, and in turn $v_i(t) = (10\text{ V})\cos 10^7 t$. Consequently, since all the envelope nformation has been stripped from $v_i(t)$, the circuit functions as an amplitude limiter rovided that no failure-to-follow distortion occurs.

xample 10.3–3 For the circuit shown in Fig. 10.3–10, make an accurate sketch of $_o(t)$, $v_o'(t)$, and $v_i(t)$.

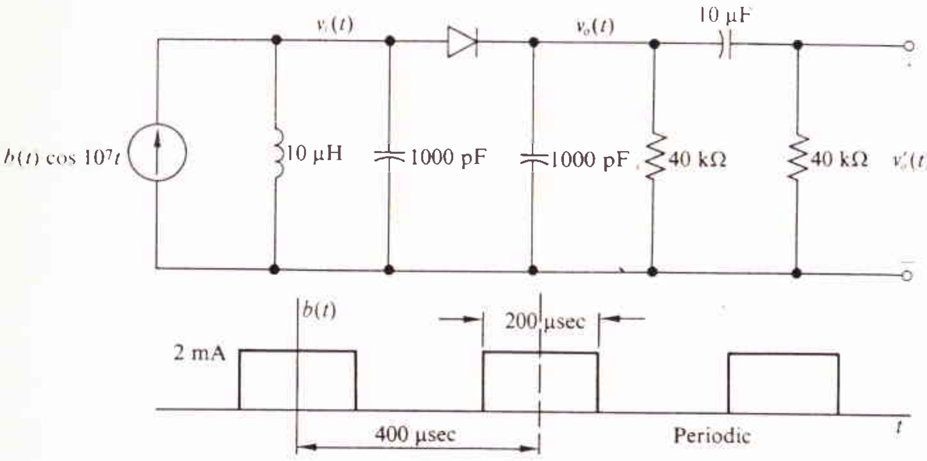


Figure 10.3–10

olution. Assuming that Q_T is high, that the ripple is small, and that there is no failure-follow distortion, we replace the circuit of Fig. 10.3–10 by its equivalent circuit own in Fig. 10.3–11. Since the 10 μF capacitor appears as a short circuit to the ac mponent of $b(t)/2$ and an open circuit to the dc component of $b(t)/2$, we decompose $t)/2$ into its ac and dc components as shown in Fig. 10.3–12 and solve for $v_o(t)$ and (t) by superposition. The dc component of $v_o(t)$ is equal to ($\frac{1}{2}$ mA) × 40 kΩ = 20 V;

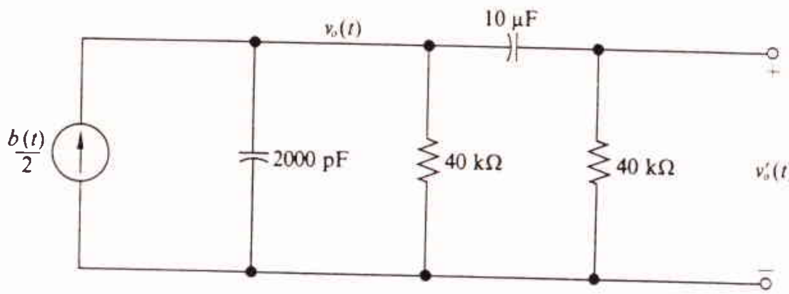


Figure 10.3–11

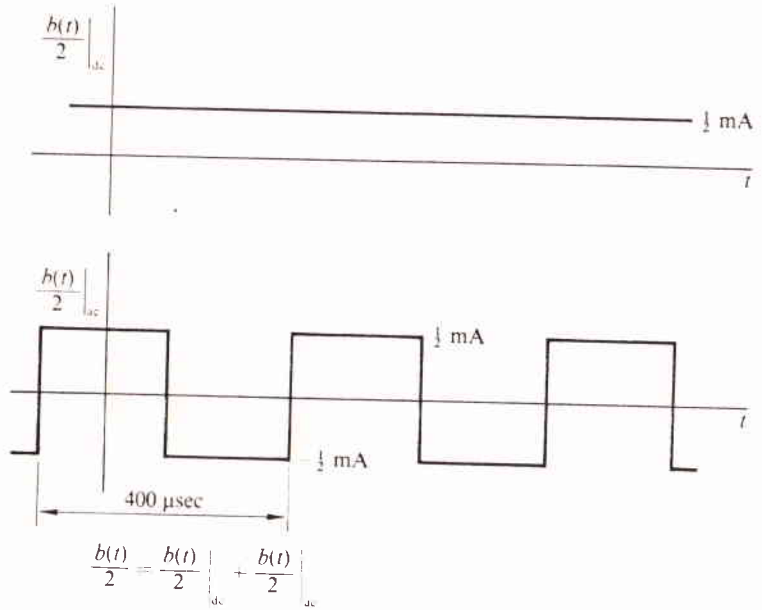


Figure 10.3–12

the dc component of $v_o'(t)$ is equal to zero. In addition, the ac components of $v_o(t)$ and $v_o'(t)$ are equal, since the 10 μF capacitor is an effective ac short. A sketch of $v_o'(t)$, $v_o(t)$, and $v_i(t) = v_o(t)\cos 10^7 t$ is shown in Fig. 10.3–13.

The presence of the coupling capacitor has significantly reduced the modulation index of $v_i(t)$ from unity. This effect is best understood by observing that the impedance presented to the modulation information [the ac components of $b(t)$] is significantly less than the impedance presented to the carrier [the dc components of $b(t)$].

For the circuit of Fig. 10.3–11, if $b(t) = I_1$, then

$$v_o(t) = V_{dc} = (40\text{ k}\Omega)(I_1/2)$$

and, in turn,

$$v_i(t) = (20\text{ k}\Omega)I_1 \cos 10^7 t.$$

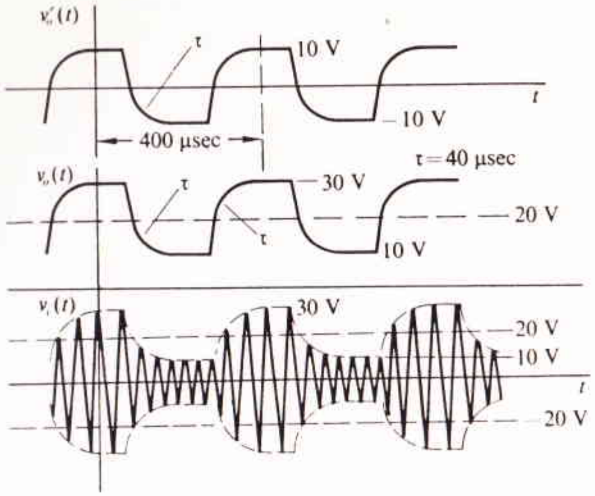


Figure 10.3–13

onsequently, $R_T = 20\,\text{k}\Omega$ and $Q_T = 200$. In addition, the fractional ripple is $\Delta V/V_{dc} = 0.0157$; hence two of our original assumptions are justified. The method or justifying the third assumption, that of no "failure-to-follow" distortion, is now onsidered.

ailure-to-Follow Distortion

n a narrowband peak envelope detector of the form shown in Fig. 10.3–1 or Fig. 0.3–6 a form of distortion known as *failure-to-follow distortion* or *diagonal clipping* may result if the circuit is improperly designed. Physically, this form of distortion esults when the input modulation on $i_i(t)$ causes the envelope of $v_i(t)$ to decrease at a ate greater than the natural decay of the parallel R_0-C_0 circuit; hence the diode fails o conduct on each cycle and $v_o(t)$ decays toward a steady-state value which is independent of the input modulation. Typical waveforms of $v_i(t)$ and $v_o(t)$ during failure-o-follow distortion are shown in Fig. 10.3–14.

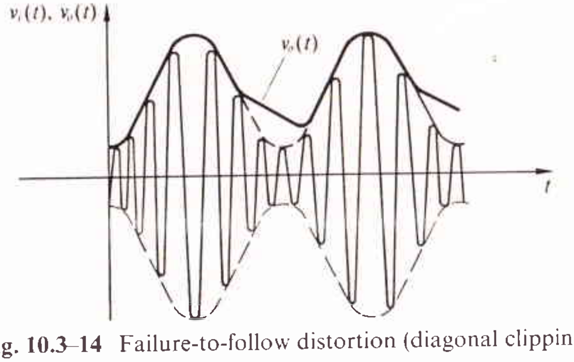


Fig. 10.3–14 Failure-to-follow distortion (diagonal clipping).

Intuitively we can make a number of observations concerning failure-to-follow distortion which are helpful in preventing its occurrence:

1. Failure-to-follow distortion always occurs on the negatively sloping envelope of the input current.
2. Failure-to-follow distortion is most likely with high modulation indices and high modulation frequencies. Both of these conditions tend to increase the negative slope of the envelopes of $i_i(t)$ and in turn $v_i(t)$, and thus increase the chance that $v_o(t)$ will not be able to follow the envelope of $v_i(t)$.
3. Failure-to-follow distortion cannot occur if the parallel RLC circuit contains no loss, that is if $R \to \infty$. If no loss is present and diode D conducts on the peak of a particular cycle of $v_i(t)$, then it must also conduct on the peak of each subsequent cycle, thus eliminating the possibility of diagonal clipping. This statement follows from the fact that the envelope of the lossless tuned circuit must increase or remain the same with D disconnected,† while with D open $v_o(t)$ decreases; hence conduction is ensured on the peak of every cycle. Thus, diagonal clipping results when the tuned circuit is sufficiently heavily loaded to permit the envelope of $v_i(t)$ to decrease more rapidly than $v_o(t)$ during the interval when D is back-biased. (It is apparent that diagonal clipping is impossible in the circuit of Example 10.3–3, since no loss exists in the tuned circuit.)
4. Failure-to-follow distortion cannot occur in the circuit of Fig. 10.3–1 if $2RC \geq R_0C_0$. Assume that for an arbitrary $t = t_i$ the diode conducts and $v_i(t_i) = v_o(t_i) = g(t_i)$. If the diode now opens, $v_o(t)$ decays exponentially to zero with an initial slope equal to $-g(t_i)/R_0C_0$. In addition, if i_i is equal to zero, the envelope of $v_i(t)$ decays exponentially to zero with an initial slope equal to $-g(t_i)/2RC$ (cf. Section 3.3). If $i_i \neq 0$, energy is supplied to the tuned circuit and the initial slope of the envelope of $v_i(t)$ is greater than $g(t_i)/2RC$; hence, if
$$\frac{g(t_i)}{2RC} \leq \frac{g(t_i)}{R_0C_0},$$
or equivalently,
$$2RC \geq R_0C_0, \tag{10.3–24}$$
the envelope of $v_i(t)$ decays less rapidly than $v_o(t)$ and the diode conducts at $t = t_i + T$. The same argument may be applied to subsequent cycles; thus the diode conducts on the peak of every cycle of $v_i(t)$ and diagonal clipping is impossible.

It is apparent at this point that, when diagonal clipping occurs, the diode current and in turn its average value drop to zero for a large number of carrier cycles; consequently, a sufficient condition to ensure that failure-to-follow distortion cannot occur is that the average diode current i_{D0} exceed zero for all time, that is,

$$i_{D0} > 0 \tag{10.3–25}$$

---

† So long as $b(t)\cos\omega_0 t$ supplies energy to the lossless tuned circuit, its voltage amplitude increases; if $b(t) = 0$, no additional energy is supplied and the voltage amplitude remains constant.

or all t. Thus by obtaining an expression for i_{D0} from the equivalent envelope deector circuits of Fig. 10.3–5 or Fig. 10.3–7 and requiring that it satisfy Eq. (10.3–25), ve obtain an inequality involving the parameters of the envelope detector which nsures that diagonal clipping cannot occur. Therefore, in addition to solving for $_o(t)$, we must solve for $i_{D0}(t)$ and check to see that it exceeds zero. If it does, the model eing used is valid. Two examples should clarify the use of this method.

xample 10.3–4 For the narrowband peak envelope detector shown in Fig. 10.3–15, letermine the relationship between m, R, C, R_0, and C_0 which ensures no diagonal lipping for all values of ω_m.

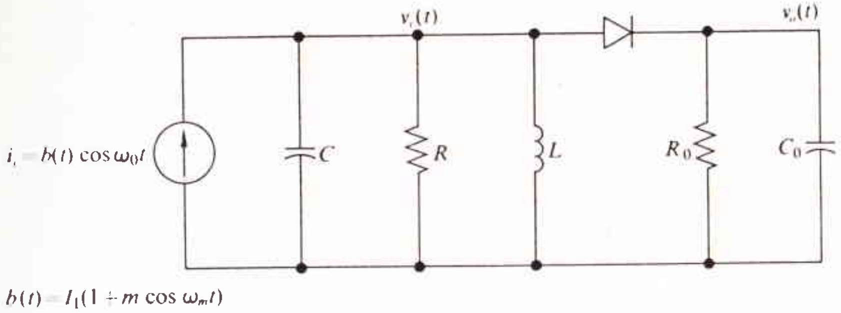


Figure 10.3–15

olution. Assuming that Q_T is high and that the output ripple is small, we replace he envelope detector by its equivalent circuit, shown in Fig. 10.3–16. For the quivalent circuit we obtain

$$i_{D0}(t) = \frac{I_1 R}{R_0 + 2R}\left[1 + m\frac{\sqrt{(\omega_m/\omega_1)^2 + 1}}{\sqrt{(\omega_m/\omega_2)^2 + 1}}\cos(\omega_m t + \phi)\right],$$

vhere

$$\omega_1 = \frac{1}{R_0 C_0}, \qquad \omega_2 = \frac{1}{(C_0 + C)(2R \| R_0)}, \qquad \text{and} \qquad \phi = \tan^{-1}\frac{\omega_m}{\omega_1} - \tan^{-1}\frac{\omega_m}{\omega_2}.$$

Clearly, for $i_{D0}(t)$ to be greater than zero for all t, we require

$$m\frac{\sqrt{(\omega_m/\omega_1)^2 + 1}}{\sqrt{(\omega_m/\omega_2)^2 + 1}} \le 1, \tag{10.3–26}$$

ince $\cos(\omega_m t + \phi)$ reaches a value of -1.

For $\omega_1 \ge \omega_2$ the coefficient of m in Eq. (10.3–26) has its maximum value of unity or $\omega_m = 0$; hence for $\omega_1 \ge \omega_2$, or equivalently for

$$R_0 C_0 \le 2RC,$$

no diagonal clipping occurs for any value of ω_m provided that $m \le 1$ (a condition which is always satisfied). This, of course, is an expected result. On the other hand,

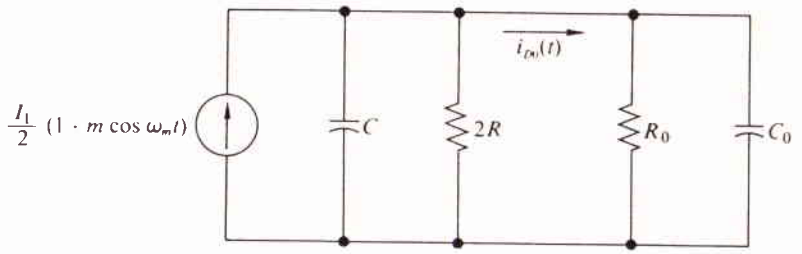


Figure 10.3–16

for $\omega_1 < \omega_2$, the coefficient of m in Eq. (10.3–26) has its maximum value of $\omega_2/\omega_1 = C_0(R_0 + 2R)\ 2R(C_0 + C)$ at $\omega_m \to \infty$; hence for $\omega_1 < \omega_2$ no diagonal clipping occurs for any value of ω_m provided that

$$m \le \frac{2R(C_0 + C)}{C_0(R_0 + 2R)}. \tag{10.3–27}$$

For example, if $2R = R_0$ and $C_0 = 3C$, then the maximum modulation index which may be used without failure-to-follow distortion is $m = \frac{2}{3}$. Also, for the case where the detector is used as a limiter and $C_0 \to \infty$,

$$m \le \frac{2R}{R_0 + 2R}. \tag{10.3–28}$$

(Note that diagonal clipping does occur in the second part of Example 10.3–2.)

Example 10.3–5 With the assumption that C_C is an ac short circuit at the frequency ω_m, determine the maximum value of modulation index m which does not cause diagonal clipping in the envelope detector of Fig. 10.3–17.

Solution. Again assuming a high Q_T and a low output ripple, we can replace the envelope detector by its equivalent circuit, shown in Fig. 10.3–18. For this circuit, $i_{D0}(t)$ is given by

$$i_{D0}(t) = \frac{I_1}{6} + \frac{mI_1}{4}\cos\omega_m t$$

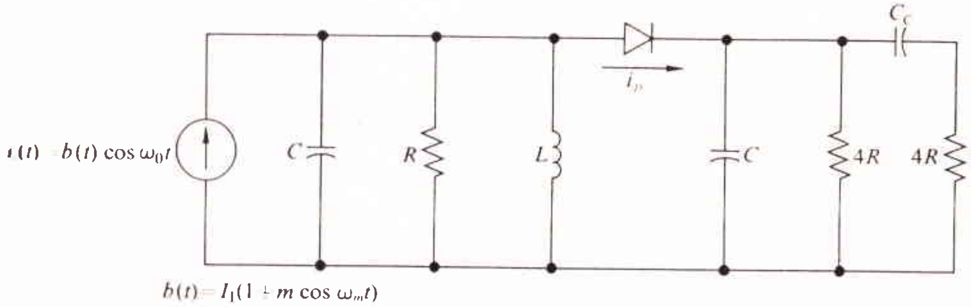

C_C
i_D
$i(t) = b(t)\cos\omega_0 t$
C
R
L
C
$4R$
$4R$
$b(t) = I_1(1 + m\cos\omega_m t)$

Figure 10.3–17

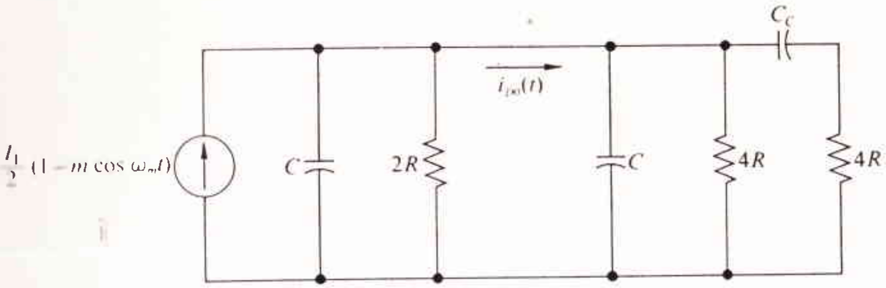


Figure 10.3–18

we assume that C_C is an ac short circuit). Hence, for $i_{p0}(t)$ to be greater than zero or all t,

$$m < \tfrac{2}{3}.$$

f the size of the output resistor or the resistor loading the tuned circuit is increased, he value of m that the circuit can accommodate increases.

ummary and Design

n the previous analyses a number of constraints on the interrelationship of parameters were required for the correct operation of the peak envelope detector shown n Fig. 10.3–1. These constraints are summarized as follows:

1) $1 \ll Q_T = \omega_0 R_T C = \dfrac{Q_{T0}Q_{TU}}{Q_{T0} + Q_{TU}}$, (selectivity)

2) $\dfrac{\tau_0}{T} = \dfrac{R_0 C_0}{T} \gg 1$, (low ripple)

3) $\omega_m \le \omega_3 = \dfrac{1}{2R_T(C + C_0)}$, (bandwidth sufficient to pass modulation information)

4) $R_0 C_0 \le 2RC$. (no failure-to-follow distortion for any m)

Here $Q_{TU} = \omega_0 RC$ is the unloaded Q_T of the tuned circuit, $Q_{T0} = \omega_0 C R_0/2$, and ω_m is the maximum modulation frequency. For circuits of the type shown in Fig. 10.3–6, only condition 4 in the above list must be modified slightly. To observe how these constraints may be employed in the design of an envelope detector, we first rewrite condition 4 in the form

$$\frac{R_0 C_0}{T} = \frac{\tau_0}{T} \le \frac{2RC\omega_0}{2\pi} = \frac{Q_{TU}}{\pi}. \tag{10.3–29}$$

To satisfy condition 2 and minimize the ripple, we wish τ_0/T to be as large as possible

however, Eq. (10.3–29) limits its value to Q_{TU}/π; hence we choose

$$\frac{\tau_0}{T} = \frac{Q_{TU}}{\pi}, \tag{10.3–30}$$

which is equivalent to $2RC = R_0 C_0$.

With the value of τ_0 given by Eq. (10.3–30), condition 3 may be rearranged in the form

$$\frac{\omega_m}{\omega_0} \le \frac{1}{2Q_{TU}}. \tag{10.3–31}$$

For proper operation we desire a small amount of ripple and thus a large value of Q_{TU}; however, from Eq. (10.3–31) we see that the maximum size of Q_{TU} is limited by the ratio of ω_m/ω_0 (broadband signals require low-Q_T circuits). Consequently the peak envelope detector may be employed only when ω_0 exceeds ω_m, usually by a factor of 100 or more. Under these conditions we choose Q_{TU} as large as possible within the limitations of Eq. (10.3–31). A large value of Q_{TU} also ensures that condition 1 will be satisfied.

Condition 1 also requires a high value of Q_{T0} which we choose, once the required value of Q_T is specified, from the relationship

$$Q_{T0} = \frac{Q_T Q_{TU}}{Q_{TU} - Q_T}. \tag{10.3–32}$$

With the arbitrary specification of one desired parameter in the circuit of Fig. 10.3–1, the relationship $\omega_0 = 1/\sqrt{LC}$ and Eqs. (10.3–30), (10.3–31), and (10.3–32) provide a set of equations for completely determining all other parameters in the envelope detector. The following example demonstrates this fact.

Example 10.3–6 The circuit of Fig. 10.3–1 has an output resistor $R_0 = 10\ \text{k}\Omega$. Choose the remaining parameters in the circuit to ensure $Q_T = 25$, minimum ripple, no failure-to-follow distortion regardless of the input modulation index, and a sufficient bandwidth to accommodate modulation frequencies up to $\omega_m = 10^6$ rad/sec. The carrier frequency is $\omega_0 = 10^8$ rad/sec.

Solution. Minimum ripple, sufficient bandwidth, and no diagonal clipping are ensured by choosing

$$\frac{\tau_0}{T} = \frac{Q_{TU}}{\pi} = \frac{\omega_0}{2\pi\omega_m} = 15.9.$$

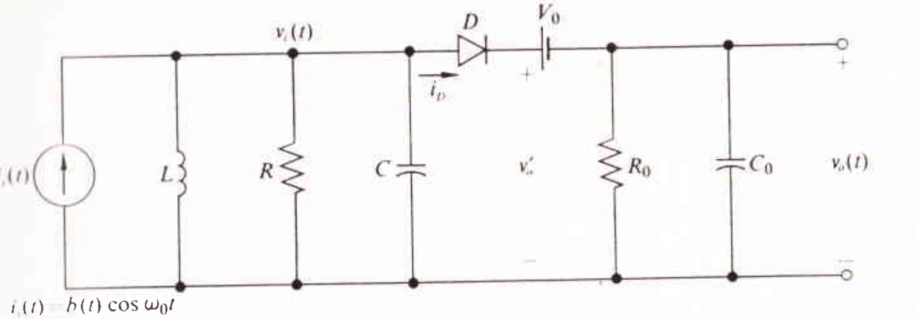
Thus $\tau_0 = R_0 C_0 = 1\ \mu\text{sec}$, $Q_{TU} = \omega_0 RC = 50$, and $RC = 0.5\ \mu\text{sec}$. Since $R_0 = 10\ \text{k}\Omega$, $C_0 = 100$ pF. By solving for Q_{T0} from Eq. (10.3–32), we obtain

$$Q_{T0} = \frac{Q_T Q_{TU}}{Q_{TU} - Q_T} = \frac{(50)(25)}{25} = 50.$$

ut $Q_{T0} = \omega_0 R_0 C/2$; therefore, $R_0 C = 10^{-6}$sec ensures $Q_T = 25$. Again, with $_0 = 10\,\text{k}\Omega$ we obtain $C = 100$ pF, and with $RC = 0.5\,\mu$sec we obtain $R = 5\,\text{k}\Omega$. inally, $1/\omega_0^2 C = 1\,\mu H = L$.

.4 PRACTICAL NARROWBAND PEAK ENVELOPE DETECTOR

a physical diode is placed in the envelope detector circuit of Fig. 10.3–1, the detector o longer functions properly for all levels of the input current $i_i(t)$. Specifically, nless the amplitude of $i_i(t)$ is sufficiently large to develop a voltage $v_i(t)$ across the uned circuit which exceeds the diode turn-on bias, then the diode remains reverse iased and $v_o = 0$. To focus attention on this problem we model the physical diode y an ideal diode in series with a battery of V_0 as shown in Fig. 10.4–1. For silicon iodes V_0 is closely approximated by 650 mV; for germanium diodes, V_0 is closer to 20 mV. From Fig. 10.4–1 it is apparent that the effect of V_0 is to keep the diode everse biased for small values of $i_i(t)$ or, equivalently, to induce failure-to-follow istortion for input modulation indices close to unity.



ig. 10.4–1 Narrowband peak envelope detector with physical diode replaced with its model.

However, if the input drive level is sufficiently high and the input modulation ndex is sufficiently low so that failure-to-follow distortion does not occur, if Q_T is high, and if the output ripple is small, then $v_o'(t)$ follows the envelope of $v_i(t)$ and $v_o(t) =$ $v_o'(t) - V_0$. That is, if $v_i(t) = g(t)\cos\omega_0 t$, then $v_o'(t) = g(t)$ and

$$v_o(t) = g(t) - V_0. \tag{10.4–1}$$

n addition, the diode current flows in sufficiently narrow pulses so that the detector an be replaced by the equivalent circuit of Fig. 10.4–2. With the aid of this circuit $v_o'(t)$ and $v_o(t)$ may be determined as functions of $b(t)$ and, in addition, $i_{D0}(t)$ may be obtained to check that V_0 has not induced failure-to-follow distortion.

For example, if $b(t) = I_1[1 + m\cos\omega_m t]\cos\omega_0 t$, then

$$v_o'(t) = \frac{2RV_0}{2R + R_0} + \frac{R_0 R I_1}{2R + R_0}\left[1 + \frac{m\cos(\omega_m t + \theta)}{\sqrt{1 + (\omega_m/\omega_3)^2}}\right] \tag{10.4–2}$$

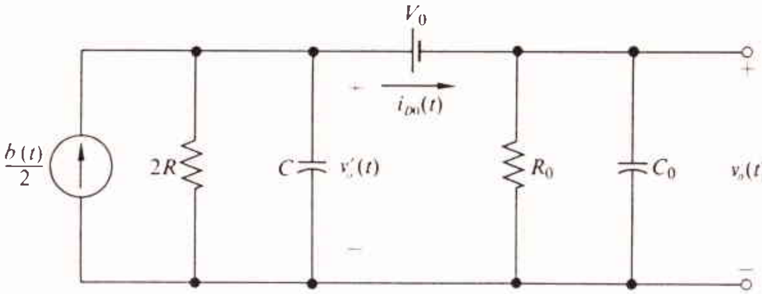


Fig. 10.4-2 Model for the circuit of Fig. 10.4–1.

and

$$v_o(t) = \frac{-V_0 R_0}{2R + R_0} + \frac{R_0 R I_1}{2R + R_0}\left[1 + \frac{m\cos(\omega_m t + \theta)}{\sqrt{1 + (\omega_m/\omega_3)^2}}\right]. \tag{10.4–3}$$

where $\theta = -\tan^{-1}(\omega_m/\omega_3)$ and $\omega_3 = 1/(R_0 \| 2R)(C + C_0)$. In addition,

$$i_{D0}(t) = -\frac{V_0}{2R + R_0} + \frac{I_1 R}{2R + R_0}\left[1 + m\frac{\sqrt{(\omega_m/\omega_1)^2 + 1}}{\sqrt{(\omega_m/\omega_3)^2 + 1}}\cos(\omega_m t + \phi)\right]. \tag{10.4–4}$$

where $\phi = \tan^{-1}(\omega_m/\omega_1) - \tan^{-1}(\omega_m/\omega_3)$ and $\omega_1 = 1/R_0 C_0$. To ensure that no failure-to-follow distortion occurs, we require $i_{D0}(t) > 0$ for all t, or equivalently,

$$1 - n > m\frac{\sqrt{(\omega_m/\omega_1)^2 + 1}}{\sqrt{(\omega_m/\omega_3)^2 + 1}}, \tag{10.4–5}$$

where $n = V_0/I_1 R$. If $2RC \geq R_0 C_0$ (the case which ensures no failure-to-follow distortion for $V_0 = 0$), then $\omega_3 \geq \omega_1$ and the coefficient of m in Eq. (10.4–5) achieves its maximum value of unity for $\omega_m = 0$. Thus to ensure no failure-to-follow distortion for all values of ω_m, we require

$$1 - n > m. \tag{10.4–6}$$

From Eq. (10.4–6) we observe that n approaches zero as V_0 is decreased, I_1 is increased, or R is increased. Consequently, for an envelope detector employing a physical diode to demodulate AM signals with modulation indices close to unity, we choose a germanium diode to keep V_0 as small as possible and we make the impedance of the tuned circuit as high as possible. (In the last section we saw that we could choose one parameter of the envelope detector arbitrarily.) In addition, as a practical matter, we choose a diode with a stray capacity much less than C or C_0. With this accomplished, the remaining components in the detector may be selected by the method discussed in Section 10.3. Here again, if the tuned circuit contains no loss, diagonal clipping is impossible regardless of the size of V_0.

10.5 BROADBAND PEAK ENVELOPE DETECTOR

An envelope detector which linearly demodulates AM signals over a broad range of input carrier frequencies is shown in Fig. 10.5–1. For this circuit the diode conducts at the peak of each carrier cycle, thus increasing the capacitor voltage $v_o(t)$ to the peak value of $v_i(t)$ less the turn-on voltage of the diode. As $v_i(t)$ decreases from its peak value, the diode is reverse biased and the capacitor holds the voltage until the diode again conducts at the peak of the next carrier cycle. The output voltage $v_o(t)$, therefore, follows the envelope of $v_i(t)$ shifted down by the turn-on voltage of the diode.

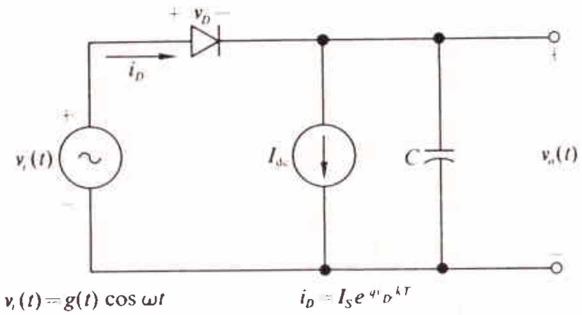


Fig. 10.5–1 Broadband peak envelope detector.

The dc current source in the network serves two valuable purposes. First, it provides a discharge path for the capacitor when the diode is reverse biased, thus permitting $v_o(t)$ to follow negatively sloping envelopes of $v_i(t)$ without failure-to-follow distortion. Secondly, it ensures that the diode turns on at the peak of each input cycle even when the input amplitude drops below the nominal turn-on voltage of the diode. Consequently, even with a silicon diode in the detector, the output faithfully responds to variations in the input envelope down to very low levels. The slightly undesirable effect of I_{dc} is the negative diode turn-on voltage which appears at the output when $v_i(t) = 0$. This voltage can readily be compensated for in circuits where its presence is undesirable.

To obtain a more quantitative expression for $v_o(t)$ in terms of $v_i(t)$, let us consider the case where $v_i(t) = V_1 \cos \omega t$ and where C is sufficiently large so that over a carrier cycle $v_o(t)$ remains constant at V_{dc}. For this case the diode current may be written as

$$\begin{aligned} i_D(t) &= I_S e^{(qV_1/kT)\cos \omega t} e^{-qV_{dc}/kT} \\ &= I_S e^{x \cos \omega t} e^{-qV_{dc}/kT} \end{aligned} \tag{10.5–1}$$

with an average value

$$\overline{i_D} = I_S I_0(x) e^{-qV_{dc}/kT}, \tag{10.5–2}$$

where $x = qV_1/kT$ and $I_0(x)$ is the modified Bessel function of order zero. Since

the steady state $\overline{i_D} = I_{dc}$, we can rearrange Eq. (10.5–2) to yield

$$\begin{aligned} V_{dc} &= \frac{kT}{q} \ln I_0(x) - \ln \frac{I_{dc}}{I_S} \\ &= \frac{kT}{q} \ln I_0\left(\frac{qV_1}{kT}\right) - V_0, \end{aligned} \tag{10.5–3}$$

where in this case $V_0 = \ln (I_{dc}/I_S)$, which has a value in the range of 200 mV for germanium diodes and 650 mV for silicon diodes. A plot of

$$f(V_1) = \frac{kT}{q} \ln I_0\left(\frac{qV_1}{kT}\right)$$

is shown in Fig. 10.5–2 for $kT/q = 26$ mV. It should be observed from Fig. 10.5–2 that $f(V_1)$ is remarkably linear for values of V_1 in excess of 50 mV and thus may be closely approximated by

$$f(V_1) = 0.997V_1 - (54 \text{ mV}) \approx V_1. \tag{10.5–4}$$

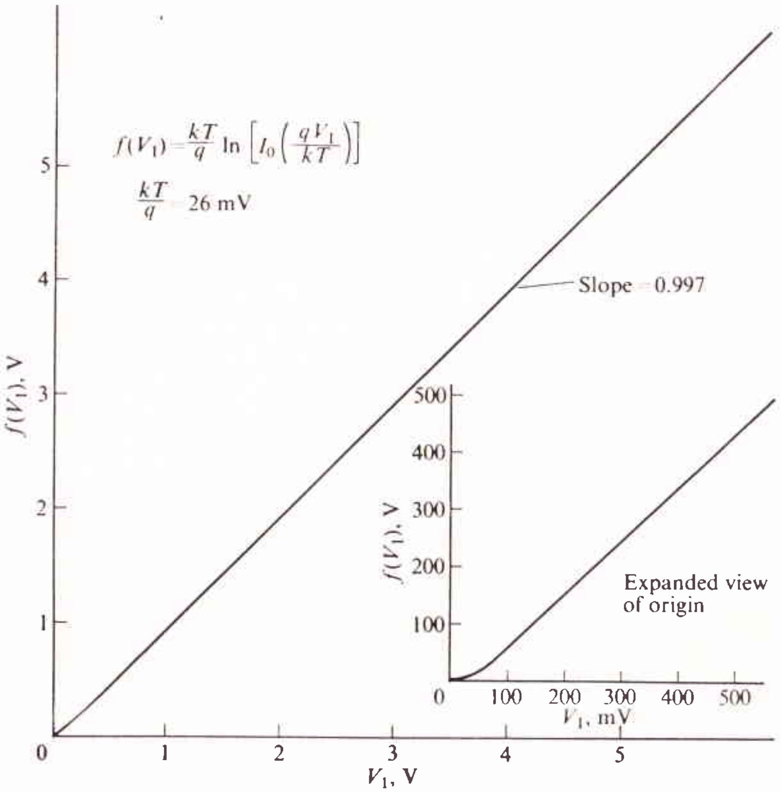


Fig. 10.5–2 Plot of $f(V_1)$ vs. V_1 with $kT/q = 26$ mV.

Consequently, with $f(V_1) \approx V_1$, the output voltage is given by

$$V_{dc} = V_1 - V_0 \tag{10.5–5}$$

for values of V_1 as low as 50 mV regardless of whether a germanium or silicon diode is being employed. Equation (10.5–5) makes it obvious that for $V_1 > 50$ mV the exponential diode in the circuit of Fig. 10.5–1 may be modeled by an ideal diode in series with a battery of value V_0, as was assumed in Section 10.4.

On the other hand, if $qV_1/kT \ll 26$ mV, then

$$I_0\left(\frac{qV_1}{kT}\right) = 1 + \tfrac{1}{4}\left(\frac{qV_1}{kT}\right)^2 + \cdots \tag{10.5–6}$$

and

$$f(V_1) \approx \frac{kT}{4q}\left(\frac{qV_1}{kT}\right)^2 = \frac{qV_1^2}{4kT}. \tag{10.5–7}$$

Therefore,

$$V_{dc} = \frac{qV_1^2}{4kT} - V_0 \tag{10.5–8}$$

and we observe that the peak envelope detector functions as a square-law detector for V_1 in the millivolt range. Unfortunately, V_0 must be removed from V_{dc} by subtraction or capacitive dc blocking before the square-law property can be fully utilized.

Output Ripple

Even if the output capacitor is large, a certain amount of ripple exists on $v_o(t)$ because of the narrow current pulses which flow each time the diode conducts to recharge the capacitor. It is apparent from Eq. (10.5–1) that the current pulses have the waveform shown in Fig. 4.5–3 which is indeed narrow for $x \gg 1$ or $V_1 \gg 26$ mV.

If we model these narrow pulses as impulses of strength TI_{dc}, where T is the period of the input carrier, then both the impulse model and the actual pulses supply the same charge to the capacitor when the diode conducts. With the impulse model, however, it is apparent that each pulse increases the capacitor voltage by

$$\Delta V = \frac{I_{dc}T}{C}. \tag{10.5–9}$$

(A slightly smaller value of ΔV is obtained if the actual current pulse shape is used.)

As should be expected, the ripple varies directly with the current I_{dc} and inversely with the capacitance and carrier frequency. Unlike the ripple in the narrowband peak envelope detector, the ripple ΔV (Eq. 10.5–9) is independent of V_{dc}. This effect is a direct result of providing I_{dc} to discharge C in lieu of a resistor.

Failure-to-Follow Distortion

Now, if $v_i(t) = g(t)\cos\omega_0 t$ and we assume that $g(t)$ remains greater than 50 mV most of the time, then we may model the exponential diode of Fig. 10.5–1 by an ideal diode in series with a battery of value V_0. If, in addition, we assume that the output ripple is small and that the diode conducts on the peak of each carrier cycle, then the output voltage v_o is given by

$$v_o(t) = g(t) - V_0, \tag{10.5–10}$$

which is the desired output voltage.

To ensure that the diode conducts on the peak of each carrier cycle it is sufficient to require, with the diode in Fig. 10.5–1 reverse biased, that the output voltage decay more rapidly than the envelope $g(t)$. Specifically, diode conduction is ensured if

$$\dot{g}(t) \geq -\frac{I_{dc}}{C} \tag{10.5–11}$$

for all t. If the inequality of Eq. (10.5–11) is not satisfied, failure-to-follow distortion of the form shown in Fig. (10.3–14) results.

The requirement of a small amount of ripple means that $I_{dc}T/C = I_{dc}/fC$ must be small in comparison with the average value of $g(t)$; on the other hand, the requirement of no failure-to-follow distortion necessitates $I_{dc}/C \geq -\dot{g}(t)$ for all t. If $g(t) = A[1 + \cos\omega_m t]$, then these two conditions require that

$$A\omega_m \leq \frac{I_{dc}}{C} \ll \frac{A\omega}{2\pi}. \tag{10.5–12}$$

Clearly, Eq. (10.5–12) can be satisfied only if $\omega \gg \omega_m$, as was also the case for the narrowband peak envelope detector.

Developing I_{dc} with Resistor Bias

If the current source in the detector of Fig. 10.5–1 is practically implemented by a larger resistor in series with a negative supply voltage V_{CC}, the circuit shown in Fig. 10.5–3 results. For this circuit, with $v_i(t) = V_1\cos\omega t$ and with a large value of C,

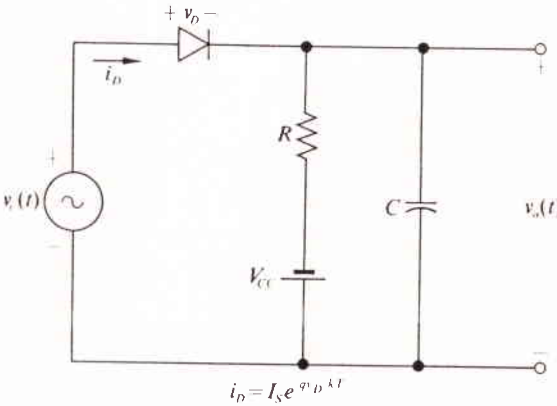


Fig. 10.5–3 Broadband peak envelope detector with resistor bias.

$v_o = V_{dc}$. Since in this case the average diode current [given by Eq. (10.5–2)] must equal $(V_{dc} + V_{CC})/R$, V_{dc} is related to $x = qV_1/kT$ by the equation

$$e^{qV_{dc}/kT}\left(\frac{V_{dc}}{V_{CC}} + 1\right) = \frac{I_S R}{V_{CC}} I_0(x). \tag{10.5–13}$$

For $V_{CC} \gg kT/q$, Eq. (10.5–13) may be closely approximated by (cf. footnote on page 173)

$$e^{qV_{dc}/kT} = \frac{I_S R}{V_{CC}} I_0(x), \tag{10.5–14}$$

from which we obtain

$$V_{dc} = \ln I_0(x) - V_0, \tag{10.5–15}$$

where $V_0 = -\ln(I_S R/V_{CC})$. In particular, the value of V_{dc} obtained from Eq. (10.5–13) differs by less than 1% from the corresponding value of V_{dc} obtained from Eq. (10.5–14) if $V_{CC} > 2.6$ V (a condition which is almost always true in practice). Consequently, by comparing Eqs. (10.5–3) and (10.5–15) we see that the output voltage is not affected by replacing the current source I_{dc} by the series R-V_{CC} combination.

On the other hand, the ripple voltage and the condition for no failure-to-follow distortion are affected by the bias change. For the circuit of Fig. 10.5–3 the ripple voltage ΔV (assuming narrow current pulses) is given by

$$\Delta V \approx \left(\frac{V_{CC} + V_{dc}}{R}\right)\frac{T}{C}; \tag{10.5–16}$$

the condition which ensures no failure-to-follow distortion is

$$\dot{g}(t) \geq -\frac{V_{CC} + g(t) - V_0}{RC}, \tag{10.5–17}$$

which must be satisfied for all t. The reader should convince himself of the validity of Eqs. (10.5–16) and (10.5–17).

Example 10.5–1 Determine the maximum value of R for the circuit shown in Fig. 10.5–4 which ensures no failure-to-follow distortion.

Solution. If failure-to-follow distortion occurs, it must occur on the negatively sloping portion of $g(t)$ between 0 and 10 μsec. During this interval,

$$g(t) = (5\text{ V})\left(1 - \frac{t}{10\ \mu\text{sec}}\right) \quad \text{and} \quad -\dot{g}(t) = \tfrac{1}{2}\text{ V}/\mu\text{sec}.$$

However, from Eq. (10.5–17), if

$$\tfrac{1}{2}\text{ V}/\mu\text{sec} \leq \frac{(15\text{ V})\left(1 - \dfrac{t}{30\ \mu\text{sec}}\right)}{R(1000\text{ pF})} \tag{10.5–18}$$

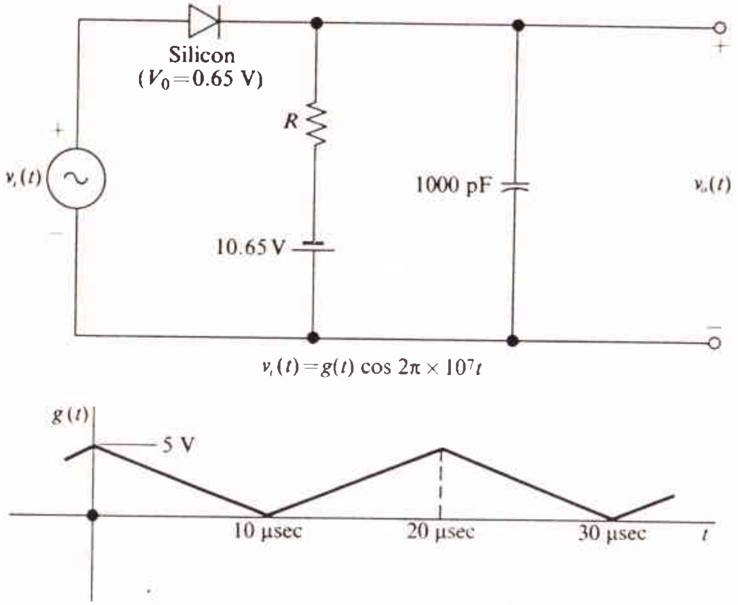


Figure 10.5–4

for $0 \leq t \leq 10\ \mu$sec, then no failure-to-follow distortion occurs. Clearly, if Eq. (10.5–18) is satisfied for $t = 10\ \mu$sec, it is satisfied over the entire interval of interest; therefore, with $t = 10\ \mu$sec Eq. (10.5–18) reduces to

$$R \leq 20\text{ k}\Omega.$$

Thus $R_{\max} = 20$ kΩ.

PROBLEMS

10.1 For the circuit shown in Fig. 10.P–1,

$$v_2(t) = \frac{g(t)}{2}\cos 10^7 t - \frac{\hat{g}(t)}{2}\sin 10^7 t,$$

where $g(t) = (1\text{ V})\cos 10^3 t$ and $v_1(t)$ has the form shown in the figure. Find $v_o(t)$.

10.2 Modify the circuit of Fig. 8.4–2 so that it functions as a synchronous detector.

10.3 For the circuit shown in Fig. 10.P–2, v_a is related to v_i by the relationship $v_a = Kv_i^2$. Find an expression for $v_o(t)$ and show that the circuit functions as a synchronous detector if $B \gg A$. Show also that with $B = 0$ the circuit functions as an envelope detector if $m \ll 1$.

10.4 The half-wave rectifier in Fig. 10.1–6 is replaced by a full-wave rectifier. Sketch the frequency spectrum of v_a for this case and show that no spectral components overlap the low-frequency spectrum so long as $\omega_0 > \omega_m$. For this case, demonstrate also that the output of the average detector is doubled.

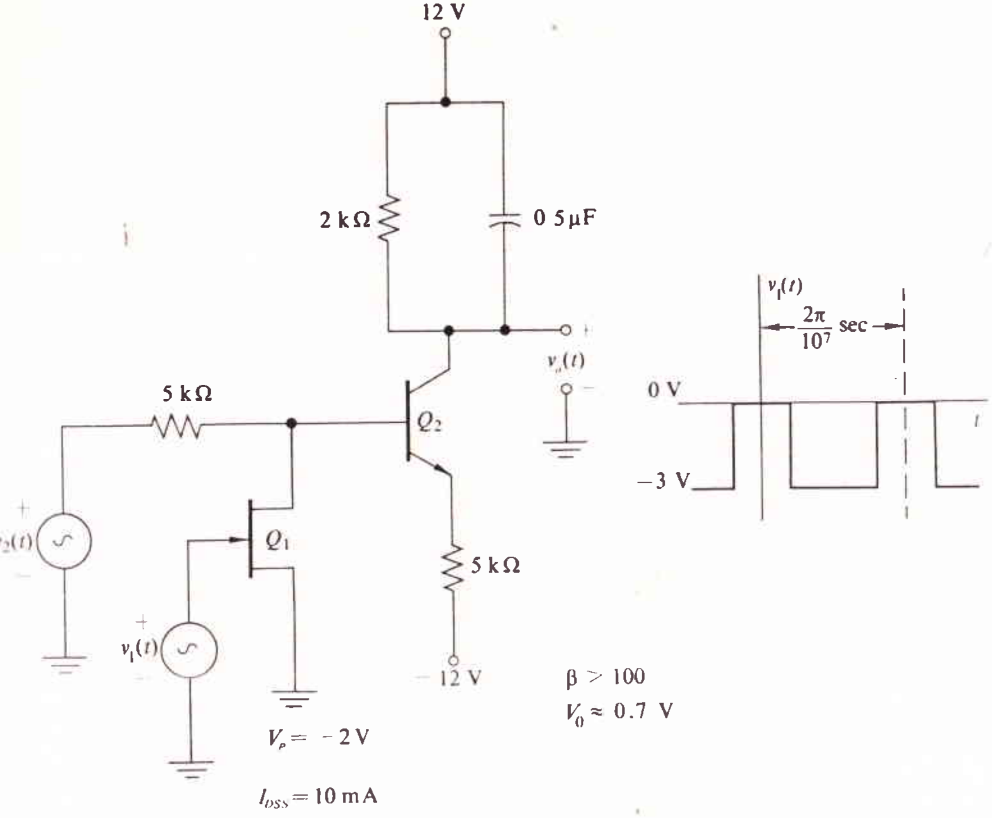


Figure 10.P–1

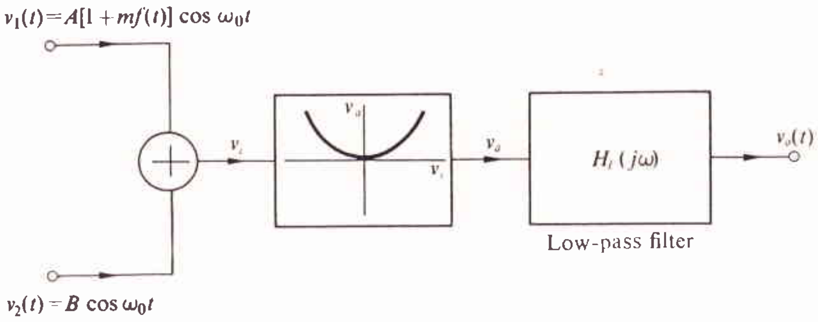

$v_1(t) = A[1 + mf(t)] \cos \omega_0 t$
$v_2(t) = B \cos \omega_0 t$
v_i
v_o
v_i
v_o
$H_l(j\omega)$
$v_o(t)$
Low-pass filter

Figure 10.P–2

10.5 The baseband signal (after frequency demodulation) of a composite SSB frequency-demodulated telephony signal has the form

$$v_r(t) = \sum_{n=1}^{8} \left[\frac{g_n(t)}{2} \cos n\omega_0 t - \frac{\hat{g}_n(t)}{2} \sin n\omega_0 t \right] + A \cos \omega_0 t,$$

where the individual voice channel $g_n(t)$ is band-limited between 200 Hz and 3 kHz and $f_0 = 4$ kHz. Draw the block diagram of a system which can be employed to extract each $g_n(t)$. Include the bandwidths and center frequencies of all filters and indicate how the reference signal for each SSB demodulator is obtained.

10.6 For the circuit of Fig. 10.2–5, $R = R_0 = 5\,\text{k}\Omega$, $C = 0.1\,\mu\text{F}$, $C_0 = 2000$ pF, $V_{CC} = 12$ V, and

$$v_i(t) = g(t) \cos 10^8 t,$$

where $g(t)$ is shown in Fig. 10.P–3. Evaluate $v_o(t)$.

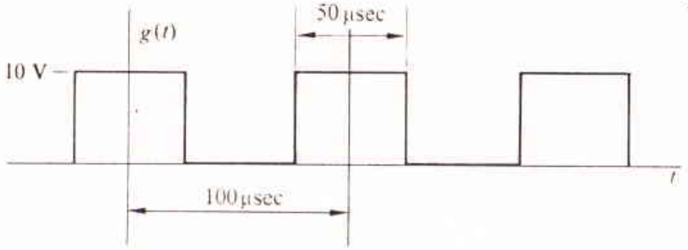


Figure 10.P–3

10.7 A "back" diode is a Zener diode which has a breakdown voltage within a few millivolts of zero. The $v_D - i_D$ characteristic for such a diode is given by

$$i_D = I_S(e^{qv_D/kT} - 1), \qquad v_D \geq 0,$$

$$i_D = v_D/r, \qquad v_D < 0,$$

where r is of the order of 10 Ω. A silicon "back" diode with $r = 20\,\Omega$ and $I_S = 10^{-16}$ A is placed in the circuit of Fig. 10.2–1 in which $R = 180\,\Omega$, $\alpha = 1$, $R_0 = 10\,\text{k}\Omega$, $C_0 = 0.01\,\mu\text{F}$, and

$$v_i(t) = V_1 \cos 10^8 t.$$

Evaluate and plot the v_o-vs.-V_1 transfer characteristic for $0 \leq V_1 < 2$ V. Over approximately what range of V_1 does the circuit function as a linear average envelope detector?

10.8 For the circuit shown in Fig. 10.P–4,

$$v_i(t) = (0.25\text{ V})(1 + \cos 10^4 t) \cos 10^7 t.$$

Find $v_o(t)$. How many decibels below the output signal is the carrier ripple?

10.9 For the circuit shown in Fig. 10.2–9, $R_1 = R_2 = 1\,\text{k}\Omega$, $R_3 = 50\,\text{k}\Omega$, $R_4 = 500\,\text{k}\Omega$, $C = 200$ pF, $A = 2000$, and

$$v_i(t) = (1\text{ V})(1 + m \cos 10^4 t) \cos 10^8 t.$$

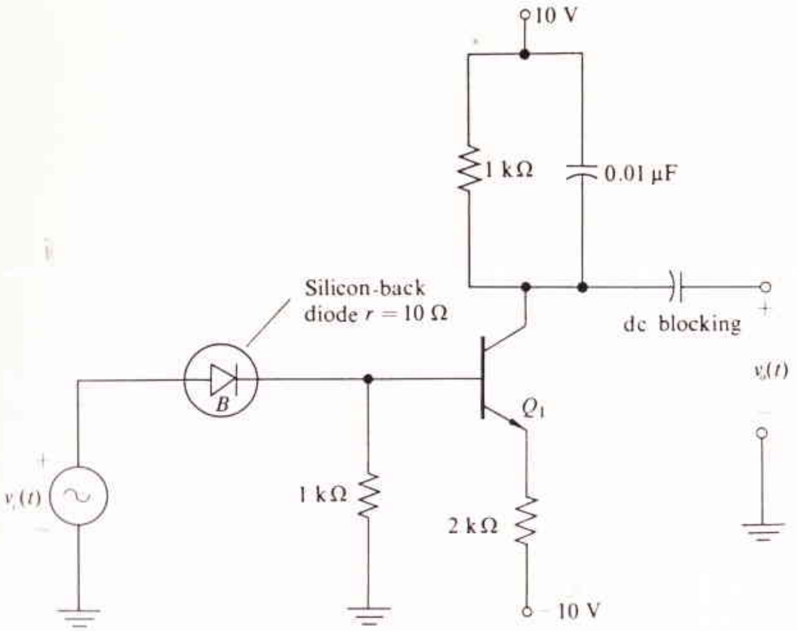


Figure 10.P–4

Find $v_o(t)$. If A is now reduced to 9 and D_1 and D_2 are hot carrier diodes with turn-on voltage $V_0 \approx 0.4$ V, determine the maximum size of m for linear demodulation.

10.10 For the circuit of Fig. 10.3–1, $R = 20\,\mathrm{k\Omega}$, $R_0 = 40\,\mathrm{k\Omega}$, $L = 10\,\mu\mathrm{H}$, $C_0 = C = 1000$ pF, and

$$i_i(t) = (10\,\mathrm{mA}) \cos 10^7 t.$$

Find $v_i(t)$, $v_o(t)$, the conduction angle of $i_D(t)$, and the fractional ripple of $v_o(t)$.

10.11 For the circuit of Fig. 10.3–1, $R = R_0/2$, $C = C_0$, and $i_i = I_1 \cos \omega_0 t$. Plot the total harmonic distortion of $v_i(t)$ vs. Q_T for the values of Q_T sufficiently large that $i_D(t)$ may be modeled as shown in Fig. 10.3–3.

10.12 For each of the circuits shown in Fig. 10.P–5, let $b(t) = I_1 = 5$ mA and determine $v_o(t$ $v_o'(t)$, $v_i(t)$, Q_T, the fractional ripple, the diode conduction angle, and the THD for $v_i(t)$.

10.13 If $b(t)$ has the form shown in Fig. 10.P–6, determine v_i, v_o, and v_o' for each of the circuits shown in Fig. 10.P–5.

10.14 Repeat Problem 10.13 with $b(t) = (5\,\mathrm{mA})(1 + \cos 2 \times 10^4 t)$.

10.15 What is the minimum value of R which may shunt L in the circuit of Fig. 10.P–5(a) without causing diagonal clipping?

10.16 What is the minimum value of R which may shunt L in the circuit of Fig. 10.P–5(b) without causing diagonal clipping if $b(t) = (5\,\mathrm{mA})(1 + \frac{1}{2} \cos 2 \times 10^4 t)$?

10.17 For the circuit shown in Fig. 10.P–7, determine expressions, for $v_1(t)$, $v_2(t)$, and $v_3($ assuming that $g(t) = (1\,\mathrm{mV})(1 + \cos 10^4 t)$.

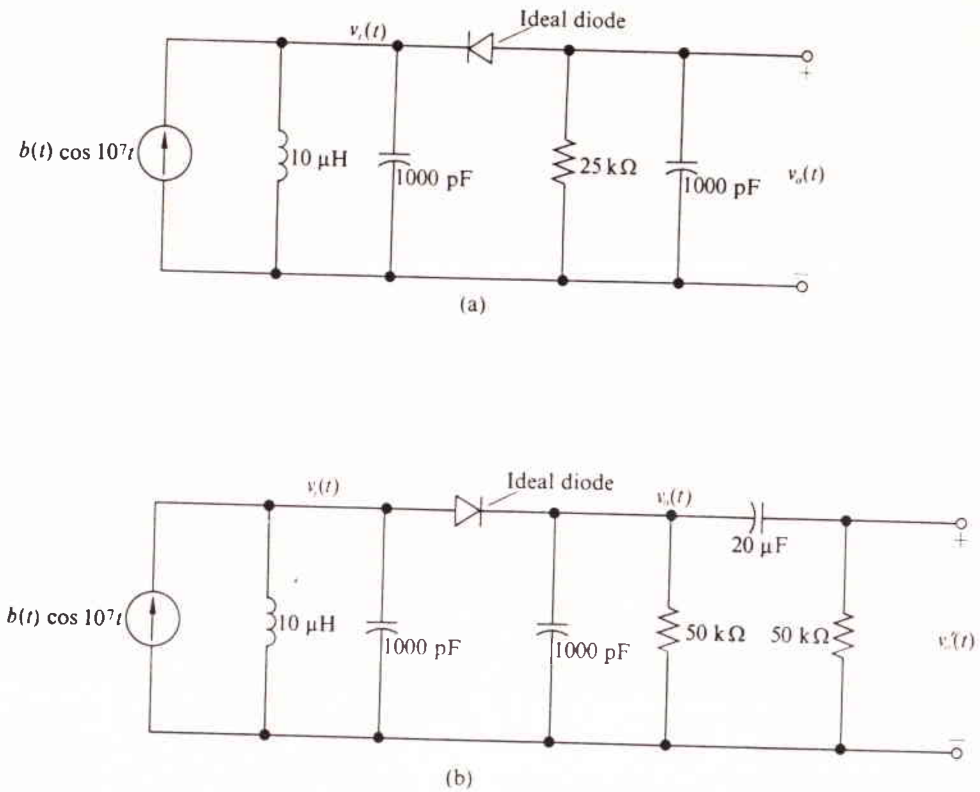


Figure 10.P–5

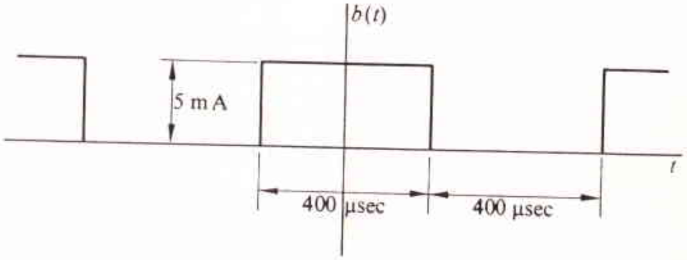


Figure 10.P–6

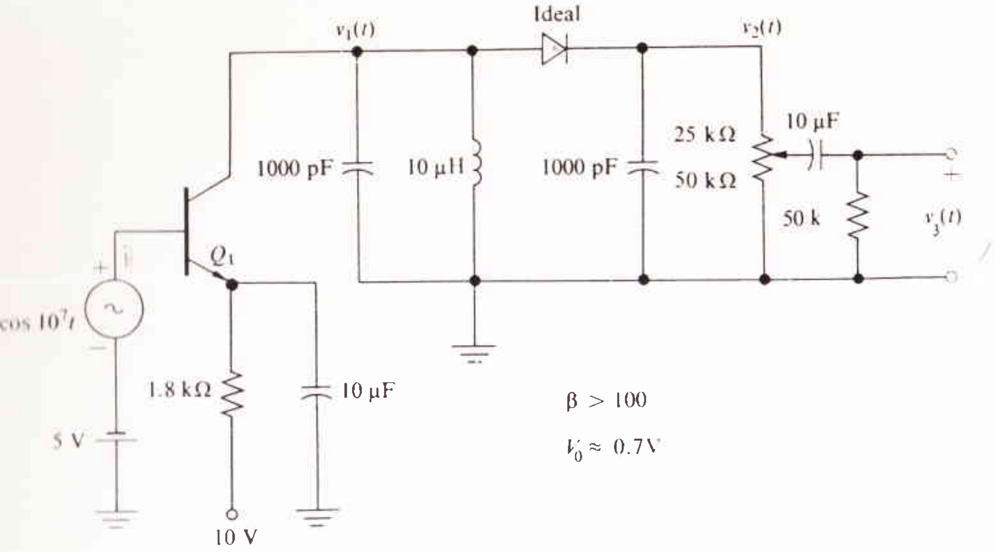


Figure 10.P–7

10.18 Repeat Problem 10.13, replacing the ideal diode with a silicon diode having $V_0 \approx 600$ mV. How much resistance can be placed across the 10 μH inductor in this case before failure-to-follow distortion occurs?

10.19 For the idealized broadband envelope detector shown in Fig. 10.P–8, show that no failure-to-follow distortion occurs if

$$-\dot{g}(t) \leq \frac{g(t)}{RC} \qquad \text{for all } t.$$

For this circuit, determine the maximum value of m which may be used without failure-to-follow distortion if

$$g(t) = A(1 + m \cos \omega_m t), \qquad \omega_m \ll \omega_0.$$

10.20 For the circuit shown in Fig. 10.5–4, $g(t) = (10 \text{ V})(1 + m \cos \omega_m t)$ and $R = 10\,\text{k}\Omega$. Find the maximum value of m (as a function of ω_m) before failure-to-follow distortion occurs.

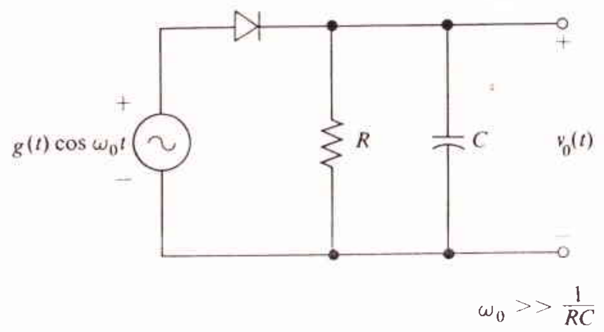


Figure 10.P–8

CHAPTER 11

GENERATION OF FM SIGNALS

In this chapter, after discussing the theoretical limitations on the transmission of FM signals through filters and nonlinear networks, we study a number of circuits for generating frequency-modulated signals. This discussion starts from the direct mechanization of the FM differential equation and proceeds to both analog and digital generation circuits.

In all cases we explore both the theoretical and the practical limits of the circuits.

11.1 FREQUENCY-MODULATED SIGNALS

If the instantaneous frequency or phase of a high-frequency sinusoid is varied in proportion to a desired low-frequency signal $f(t)$, a constant amplitude-modulated signal is generated which has its spectrum concentrated about the frequency of the unmodulated sinusoid. Thus a frequency- or phase-modulated signal, like an AM signal, can be efficiently transmitted with reasonably small antennas if the carrier frequency is sufficiently high. In addition such a signal can be frequency division multiplexed.

However, frequency- and phase-modulated signals, unlike amplitude-modulated signals, have the advantage that no information is carried in their envelope; thus atmospheric and receiver noise which introduces unwanted perturbations on the envelope of a received signal can be removed by limiting the signal amplitude before the FM or PM signal is demodulated. As a consequence, the signal-to-noise ratio at the output of an FM or PM receiver is higher than that of a comparable AM receiver which receives a signal with the same average carrier power as the received FM or PM signal and which has been subjected to the same unwanted disturbances.

The price that must be paid for the increased output signal-to-noise ratio in FM or PM receivers is increased transmitted signal bandwidth. For FM signals, the output signal-to-noise ratio increases only when the ratio of the bandwidth of the FM signal to the bandwidth of $f(t)$ is increased.

A typical phase-modulated signal has the form

$$v(t) = A \cos \phi(t) = A \cos [\omega_0 t + \Delta\phi f(t)], \tag{11.1–1}$$

where $\phi(t) = \omega_0 t + \Delta\phi f(t)$ is the instantaneous phase, A is the carrier amplitude, ω_0 is the carrier frequency, $\Delta\phi$ is the phase deviation, and $f(t)$ is a signal proportional to the modulation information with the properties

$$|f(t)|_{\max} = 1 \qquad \text{and} \qquad \overline{f(t)} = 0. \tag{11.1–2}$$

Clearly the instantaneous phase varies in direct proportion to $f(t)$ with the constant of proportionality $\Delta\phi$.

Any time-varying instantaneous phase $\phi(t)$ has associated with it an *instantaneous frequency* $\omega_i(t)$ which *by definition* is the derivative of the instantaneous phase; that is,

$$\omega_i(t) = \frac{d\phi(t)}{dt}. \qquad (11.1\text{–}3)$$

If, instead of varying $\phi(t)$ in proportion to $f(t)$, we vary $\omega_i(t)$ in proportion to $f(t)$, that is, if

$$\omega_i(t) = \omega_0 + \Delta\omega f(t), \qquad (11.1\text{–}4)$$

then

$$\phi(t) = \omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \theta_0 \qquad (11.1\text{–}5)$$

and $v(t)$ takes the form of a typical *frequency-modulated* signal:

$$v(t) = A \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \theta_0\right], \qquad (11.1\text{–}6)$$

where $\Delta\omega$, the frequency deviation, is the constant of proportionality relating $f(t)$ to the instantaneous frequency and θ_0 is an arbitrary constant phase which may be taken as zero without loss of generality. A plot of $v(t)$ given by Eq. (11.1–6) with a sawtooth signal for $f(t)$ is shown in Fig. 11.1–1.

It is apparent from Eq. (11.1–3) that we can convert any PM signal into an equivalent FM signal; therefore, by developing the properties of the FM signal, we are simultaneously developing the properties of PM signals. Hence we shall restrict our attention to the FM signal of the form given by Eq. (11.1–6).

Before continuing further we should note that there are *three* frequencies associated with an FM signal. One frequency is obviously the carrier frequency ω_0. The second frequency is the frequency deviation $\Delta\omega$, which is a measure of *how far* the instantaneous frequency departs from the carrier frequency as $f(t)$ varies between plus and minus one. The third frequency is the maximum modulation frequency ω_m [the frequency to which $f(t)$ is band-limited]. This maximum modulation frequency is a measure of *how rapidly* the instantaneous frequency varies about ω_0.

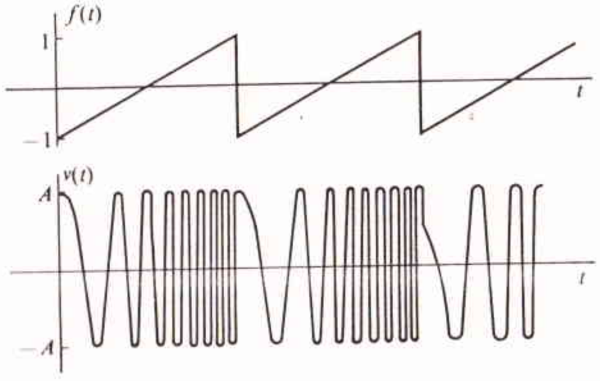


Fig. 11.1–1 Sawtooth-modulated FM signal.

We may express these three frequencies in the form of two parameters by normalizing ω_0 and ω_m to $\Delta\omega$ with the relationships

$$\beta \equiv \frac{\Delta\omega}{\omega_m} \quad \text{and} \quad D \equiv \frac{\Delta\omega}{\omega_0}, \qquad (11.1\text{–}7)$$

where β is referred to as the modulation index and D is referred to as the deviation ratio (β can have values in the thousands, whereas D is always required to be less than unity and in many practical systems may be 0.005 or even smaller).

Frequency Spectrum

To obtain the frequency spectrum of an FM signal we may, in principle, obtain the Fourier transform $V(\omega)$ of $v(t)$ given by Eq. (11.1–6). Unfortunately such an operation is not mathematically tractable for an arbitrary $f(t)$. Therefore, we restrict our initial attention to the special case of obtaining the spectrum of $v(t)$ with $f(t) = \cos\omega_m t$. Fortunately, as we shall demonstrate, the bandwidth of $v(t)$ obtained with $f(t) = \cos\omega_m t$ provides a conservative bound for the bandwidth of $v(t)$ which is modulated in frequency by any $f(t)$ band-limited to ω_m.

With $f(t) = \cos\omega_m t$, Eq. (11.1–6) may be written in the form

$$\begin{aligned} v(t) &= A\cos(\omega_0 t + \beta\sin\omega_m t). \\ &= A\cos\omega_0 t\cos(\beta\sin\omega_m t) - A\sin\omega_0 t\sin(\beta\sin\omega_m t), \qquad (11.1\text{–}8) \end{aligned}$$

which is the superposition of two AM waves, the first modulated by $\cos(\beta\sin\omega_m t)$ and the second modulated by $\sin(\beta\sin\omega_m t)$. Thus if we obtain the spectra of the two low-frequency, periodic modulating functions, we may shift them up in frequency by ω_0 to obtain the spectrum of $v(t)$.

Now $\cos(\beta\sin\omega_m t)$ and $\sin(\beta\sin\omega_m t)$ may both be expanded directly as Fourier series whose coefficients are ordinary Bessel functions of the first kind with argument β[†]:

$$\cos(\beta\sin\omega_m t) = J_0(\beta) + 2\sum_{n=1}^{\infty} J_{2n}(\beta)\cos 2n\omega_m t, \qquad (11.1\text{–}9)$$

$$\sin(\beta\sin\omega_m t) = 2\sum_{n=0}^{\infty} J_{2n+1}(\beta)\sin(2n+1)\omega_m t. \qquad (11.1\text{–}10)$$

A plot of $J_n(\beta)$ vs. β for several typical values of β and n is given in Fig. 11.1–2.

[†] Equations 9.1–42 and 9.1–43 on p. 361 of *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun, Eds., National Bureau of Standards, Applied Mathematics Series No. 55, Government Printing Office, Washington, D.C. (1964). Revised by Dover, New York (1965).

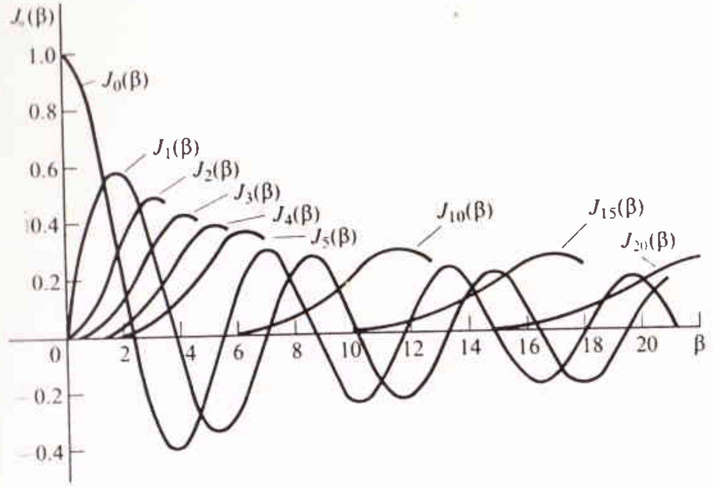


Fig. 11.1–2 Plot of $J_n(\beta)$ vs. β.

If we now recombine Eqs. (11.1–9) and (11.1–10) with Eq. (11.1–8) by making use the trigonometric identities

$$\cos x \cos y = \tfrac{1}{2}[\cos(x + y) + \cos(x - y)]$$

nd

$$-\sin x \sin y = \tfrac{1}{2}[\cos(x + y) - \cos(x - y)],$$

e finally obtain

$$\begin{aligned} v(t) = A\{&J_0(\beta)\cos\omega_0 t + J_1(\beta)[\cos(\omega_0 + \omega_m)t - \cos(\omega_0 - \omega_m)t] \\ &+ J_2(\beta)[\cos(\omega_0 + 2\omega_m)t + \cos(\omega_0 - 2\omega_m)t] \\ &+ J_3(\beta)[\cos(\omega_0 + 3\omega_m)t - \cos(\omega_0 - 3\omega_m)t] \\ &+ J_4(\beta)[\cos(\omega_0 + 4\omega_m)t + \cos(\omega_0 - 4\omega_m)t] \\ &+ \cdots\}. \end{aligned} \qquad (11.1\text{–}11)$$

rom Eq. (11.1–11) it becomes apparent that the spectrum of a sinusoidally modulated M wave contains an infinite number of sidebands occurring at frequencies ω_m, ω_m, $3\omega_m$, ... on each side of the carrier frequency ω_0. Fortunately, as can be seen rom Fig. 11.1–2, for any given modulation index β, only a finite number of Bessel unctions have values significantly different from zero and thus only the corresponding umber of upper and lower sidebands play a significant role in determining the FM pectrum. For example, with $\beta = 1$, $J_0(1) = 0.7652$, $J_1(1) = 0.4401$, $J_2(1) = 0.1149$, $J_3(1) = 0.01956$, $J_4(1) = 0.002477$, $J_5(1) = 0.0002498$, and $J_6(1) = 0.00002094$. Thus f we retain only those sidebands whose coefficient $J_n(1)$ is greater than 0.01 (i.e. hose sidebands which are greater than 1 % of the unmodulated carrier), we observe

that the FM signal with $\beta = 1$ occupies a band of frequencies of $6\omega_m$. A plot of the magnitude of the line spectrum of $v(t)$ [given by Eq. (11.1–11)] with $\beta = 1$ is shown in Fig. 11.1–3(b). Also shown in the figure are similar plots for $\beta = 0.2$, $\beta = 2.4048$, $\beta = 5$, and $\beta = 10$. In each case only those sidebands greater than 1 % of the unmodulated carrier are retained as being significant.† As expected from Fig. 11.1–2, the number of significant sidebands increases with β.

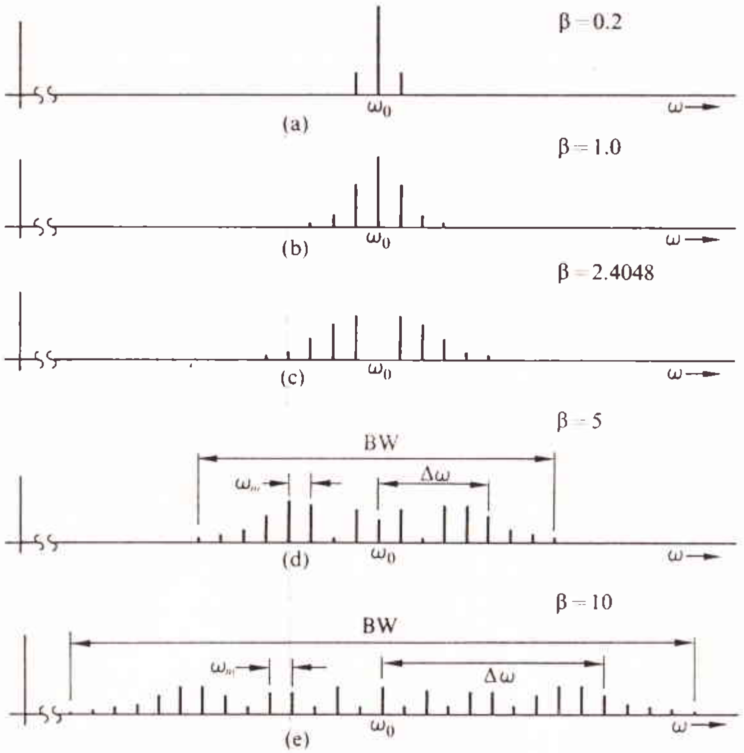


Fig. 11.1–3 Plot of amplitude spectrum of an FM signal for several values of modulation index.

From Fig. 11.1–2 we also observe that the coefficients of the various sidebands oscillate as the modulation index is increased. In particular, as β is increased from zero, $J_0(\beta)$ and thus the carrier amplitude decrease from unity. The carrier amplitude reaches zero for $\beta = 2.4048$; it then becomes negative and returns to zero for $\beta = 5.5200$; and so on. Table 11.1–1 tabulates the values of β for which $J_n(\beta) = 0$ for $n = 0, 1, 2, 3$ and 4.

Returning to Fig. 11.1–3, we observe that we can determine the bandwidth (BW) of a sinusoidally modulated FM carrier for any modulation index β by determining the number of Bessel functions $J_n(\beta)$ which exceed 0.01 and then multiplying

---

† Figure 11.1–3 would remain essentially unchanged if we retained all sidebands for which $J_n(\beta) \geq 0.005$ because $J_n(\beta)$ decreases very rapidly with increasing n.

Table 11.1-1 Values of β for which $J_n(\beta) = 0$

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$
First zero	2.405	3.832	5.136	6.380	7.588
Second zero	5.520	7.016	8.417	9.761	11.065
Third zero	8.654	10.173	11.620	13.015	14.372
Fourth zero	11.792	13.324	14.796	16.223	17.616
Fifth zero	14.931	16.471	17.960	19.409	20.827

The fact that the carrier amplitude and sideband amplitude vanish for known values of β ermits the rapid determination of the frequency deviation vs. modulation amplitude for an FM enerator whose calibration is not known. Specifically, modulation of the form $V_1 \cos \omega_m t$ is pplied to the modulator with a known value of ω_m. The amplitude of V_1 is then increased from ero until the carrier at the modulator output disappears. (The carrier is monitored on a narrow-and receiver, or on a wave or spectrum analyzer.) For this value of V_1, $\Delta\omega = 2.405\omega_m$ and a oint on the $V_1 - \Delta\omega$ calibration curve is determined. Other points may be determined from ubsequent zero values of the carrier or other sidebands.

y $2\omega_m$. We can present this bandwidth in a universal form by normalizing BW to $\Delta\omega$ and plotting it vs. β. For example, with $\beta = 1$, BW $= 6\omega_m$ and thus

$$\frac{\text{BW}}{2\,\Delta\omega} = \frac{3\omega_m}{\Delta\omega} = \frac{3}{\beta} = 3.$$

ther points obtained in a similar fashion are presented in Fig. 11.1–4, in which a mooth curve is drawn as a best fit through all the points.

For very large values of β the curve approaches an asymptote of 1. This asymptote s intuitively reasonable for any FM signal with any $f(t)$, since for a finite frequency eviation a large value of β implies a small value of ω_m or a very slowly varying $f(t)$. ery slow variations in $f(t)$ cause $v(t)$ to vary almost statically between $\omega_0 - \Delta\omega$ and $\omega_0 + \Delta\omega$; thus a filter with a bandwidth $2\,\Delta\omega$ centered about ω_0 has the minimum bandwidth to pass the slowly varying carrier of $v(t)$ undistorted. Consequently the spectrum of $v(t)$ must have BW $= 2\,\Delta\omega$ or equivalently BW/$2\,\Delta\omega = 1$.

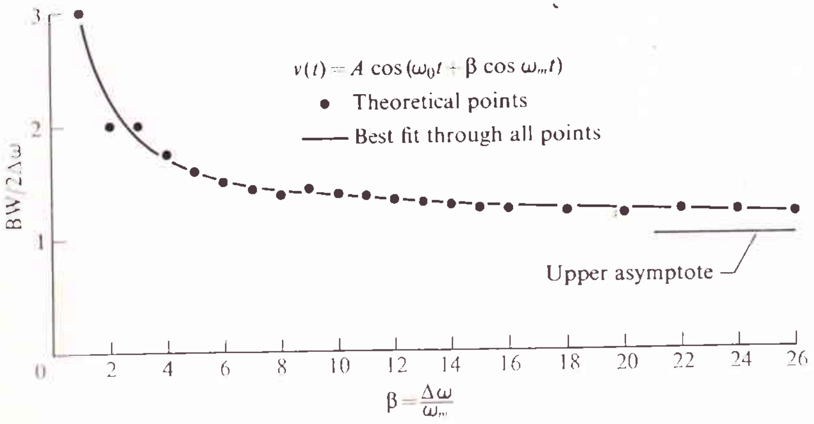


Fig. 11.1-4 Plot of normalized FM bandwidth (BW) vs. modulation index $\beta = \Delta\omega/\omega_m$.

As an example of the usefulness of the normalized bandwidth curve, consider an FM signal with a modulation frequency $\omega_m = 2\pi \times 10$ kHz and a frequency deviation of $\Delta\omega = 2\pi \times 50$ kHz. For this signal, $\beta = 5$ and thus from Fig. 11.1–4 BW/$2\,\Delta\omega = 1.6$. Consequently a bandwidth of $2\pi \times 160$ kHz is required for this sinusoidally modulated FM signal.

Similarly with $\omega_m = 2\pi \times 5$ kHz and $\Delta\omega = 2\pi \times 50$ kHz, $\beta = 10$, BW/$2\,\Delta\omega =$ 1.4, and the bandwidth is $2\pi \times 140$ kHz. As expected, a smaller bandwidth is required for a fixed frequency deviation if the modulation frequency is decreased. Consequently, for any arbitrary $f(t)$ band-limited to ω_m, the frequency components of $f(t)$ closest to ω_m have a greater effect on increasing the bandwidth of the FM signal than do those components close to $\omega = 0$. Thus if we were to model $f(t)$ as $\cos \omega_m t$, a signal with *all* of its energy concentrated at ω_m, the bandwidth of the FM signal would be greater than the corresponding bandwidth with any other $f(t)$ whose energy is distributed in the frequency range between $\omega = 0$ and $\omega = \omega_m$. As a result, the bandwidth curve of Fig. 11.1–4 provides a conservative bound for any FM signal whose modulation is band-limited to ω_m. Actually, in many practical cases, the highest-frequency sine-wave approximation gives much too conservative an estimate of bandwidth.

For high-β cases, where the peak deviation is high in comparison to the maximum frequency component of the modulating signal, there is a rule known as "Woodward's Theorem."† According to this rule, the shape of the envelope of the FM spectrum is approximately that of the amplitude probability density of the modulating function. The scale factor of this spectrum along the frequency axis is, of course, proportional to $\Delta\omega$. Intuitively this means that a high-β square-wave modulation has two impulses in frequency at $f_0 \pm \Delta f$, that a high-β triangular-wave modulation has a flat spectrum, and that a high-β Gaussian or noiselike function has a Gaussian-shaped spectrum.

Thus a combination of a number of voice channels that will combine to provide a noiselike signal produces an FM spectrum (provided that the β is large) that concentrates its energy near the carrier rather than near the edge of the band. In such cases the actual bandwidths required of filters and other circuits are substantially smaller than might be expected from the sine-wave approximation.

11.2 TRANSMISSION OF FM SIGNALS THROUGH NONLINEAR NETWORKS

In this section we shall demonstrate that an FM signal of the form

$$v(t) = V_1 \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta\right], \qquad (11.2\text{–}1)$$

---

† P. M. Woodward, "The Spectrum of Random Frequency Modulation," Telecommunications Research Establishment, Great Malvern, Worcs., England, Memo 666, December 1952.

hen passed through a nonlinear nonmemory circuit, behaves as if it were a single-equency sinusoid. Specifically, we shall show that the output $v_o(t)$ of the nonlinear rcuit may be expanded in a "Fourier-like" series of FM waves of the form

$$v_o(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos\left[n\omega_0 t + n\,\Delta\omega \int^t f(\theta)\,d\theta\right], \tag{11.2–2}$$

here the coefficients are exactly the ones that would be obtained if the circuit were riven by $V_1 \cos \omega_0 t$. This "Fourier-like" expansion makes evident the fact that a ignal proportional to the original FM signal with the same instantaneous frequency an be obtained by filtering the components of $v_o(t)$ which are centered in frequency bout ω_0. If the nonlinear circuit causes a_1 to be independent of V_1, the output of he narrowband filter is independent of V_1 and an effective FM limiter results.

On the other hand, if a narrowband filter is employed to extract the components of $v_o(t)$ centered in frequency about $n\omega_0$, an FM signal is obtained with its instantaneous frequency proportional to $f(t)$ but with n times the frequency deviation and times the carrier frequency of $v(t)$. This technique for increasing the deviation of an M wave finds wide application when a specified frequency deviation is required and et the maximum deviation available from a particular modulator produces less than he specified value. If multiplication of the carrier frequency along with the deviation s not desired, the carrier may be heterodyned back to the desired frequency by any of the techniques of Chapter 7 after the required deviation multiplication has been chieved.

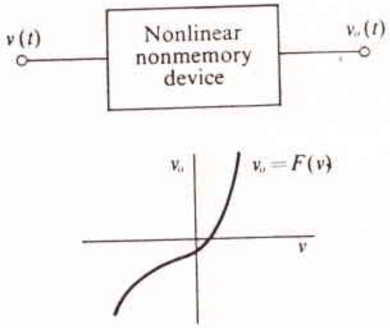


Fig. 11.2–1 Nonlinear nonmemory device characteristic.

To demonstrate the validity of Eq. (11.2–2), let us assume that $v(t)$ given by Eq. (11.2–1) is placed through a nonlinear nonmemory device, as shown in Fig. 11.2–1. Since $v(t)$ is, in general, not periodic in time, $v_o = F(v)$ is not periodic in time and a conventional Fourier series expansion of $v_o(t)$ is not possible. However, $v(t)$ is periodic with period $T = 2\pi/\omega_0$ in $\tau(t)$, where

$$\tau(t) = t + \frac{\Delta\omega}{\omega_0}\int^t f(\theta)\,d\theta, \tag{11.2–3}$$

since, as a function of τ, $v(t) = v_1(\tau)$ is given by

$$v(t) = v_1(\tau) = V_1 \cos \omega_0 \tau. \tag{11.2–4}$$

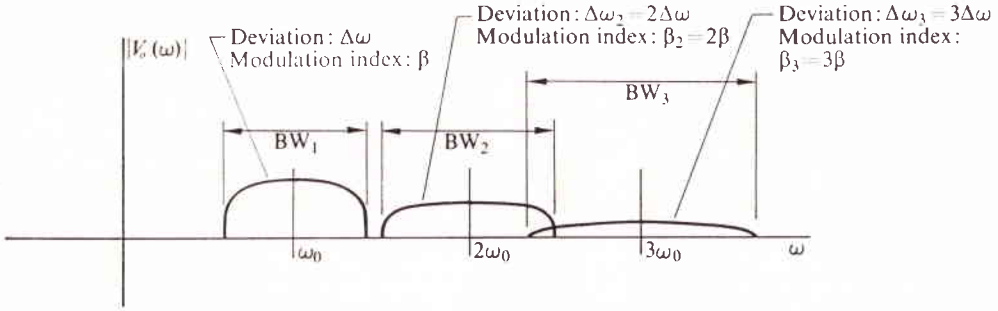
Therefore, $v_o(t) = v_{o1}(\tau) = F[v_1(\tau)]$ is a periodic function in τ and may be expanded in a Fourier series in τ of the form

$$v_o(t) = v_{o1}(\tau) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 \tau. \tag{11.2–5}$$

If τ, given by Eq. (11.2–3), is resubstituted into Eq. (11.2–3), the Fourier-like series of Eq. (11.2–2) follows immediately.

Note that the coefficients a_n in Eqs. (11.2–5) and (11.2–2) are the coefficients of the Fourier series of the output of the nonlinear device when the input is driven by a cosinusoidal signal of frequency ω_0 and amplitude V_1. Consequently all of the results of Chapters 1, 4, and 5 obtained for cosinusoidal inputs to nonlinear devices apply directly to FM signals.

If at the output of the nonlinear device we wish to reclaim an FM signal proportional to the input FM signal, we must be able to separate the fundamental FM component of $v_o(t)$ from all the higher-order FM components of $v_o(t)$. Figure 11.2–2 illustrates a typical frequency spectrum for $v_o(t)$ in which $V_0(\omega)$ is the Fourier transform of $v_o(t)$. It is apparent from the figure that the fundamental FM component of $v_o(t)$ may be extracted by filtering provided that



$|V_o(\omega)|$
Deviation: $\Delta\omega$
Modulation index: β
Deviation: $\Delta\omega_2 = 2\Delta\omega$
Modulation index: $\beta_2 = 2\beta$
Deviation: $\Delta\omega_3 = 3\Delta\omega$
Modulation index: $\beta_3 = 3\beta$
BW_1
BW_2
BW_3
ω_0
$2\omega_0$
$3\omega_0$
ω

Fig. 11.2–2 Frequency spectrum of $v_o(t)$.

$$\frac{BW_1}{2} + \frac{BW_2}{2} \le \omega_0, \tag{11.2–6}$$

where BW_1 is the total bandwidth of the fundamental FM component of $v_o(t)$ and BW_2 is the total bandwidth of the second-harmonic FM component of $v_o(t)$. By noting that $\Delta\omega_2 = 2\,\Delta\omega$, we may express Eq. (11.2–6) in the equivalent form

$$D = \frac{\Delta\omega}{\omega_0} \le \frac{1}{\dfrac{BW_1}{2\,\Delta\omega} + 2\dfrac{BW_2}{2\,\Delta\omega_2}}, \tag{11.2–7}$$

where D is the deviation ratio of the input FM signal.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
[illegible] DOCUMENTACION

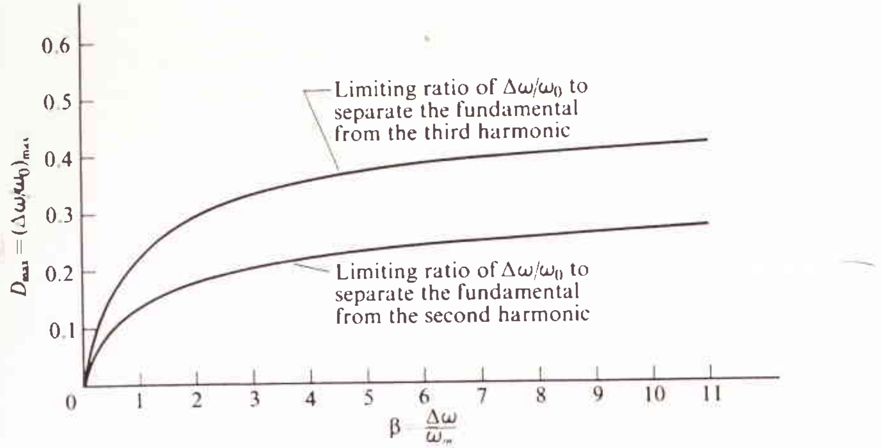


Fig. 11.2–3 Plot of the maximum deviation ratio $\Delta\omega/\omega_0$ which permits the separation of the fundamental of an FM wave from the second or third harmonic.

It is apparent that for any given value of β a conservative estimate of $\text{BW}_1/2\,\Delta\omega$ and $\text{BW}_2/2\,\Delta\omega_2$ may be obtained from Fig. 11.1–4 and thus $D_{\max}$ may be obtained as a function of β. For example, if $\beta = 5$, then $\beta_2 = 10$; hence from Fig. 11.1–4 $\text{BW}_1/2\Delta\omega = 1.6$, $\text{BW}_2/2\Delta\omega_2 = 1.4$, and $D_{\max} = 0.227$. Similarly, as $\beta \to \infty$, $\text{BW}_1/2\,\Delta\omega \to \text{BW}_2/2\,\Delta\omega_2 \to 1$ and $D_{\max} \to 0.333$. A plot of $D_{\max}$ vs. β appears as the lower curve in Fig. 11.2–3.

With the aid of the curve of Fig. 11.2–3 we can determine whether an FM signal may be obtained from a nonlinear device output by filtering. Specifically, with a knowledge of $\Delta\omega/\omega_0$ and β for a given FM signal, we plot a point on the coordinates of Fig. 11.2–3. If this point lies below the curve, an FM signal directly proportional to the input FM signal may be obtained by placing the nonlinear device output through a bandpass filter.

Figure 11.2–3 also contains a curve of $D_{\max}$ for the case where a second-harmonic FM component does not exist at the nonlinear device output. In this case the requirement for recovering the fundamental FM signal by filtering is that the spectra of the fundamental and third-harmonic signals do not overlap. In a similar fashion one may obtain $D_{\max}$ vs. β curves when one wishes to extract the nth harmonic of $v_o(t)$ by filtering. In all cases, however, if operation is close to the $D_{\max}$-curve, an extremely complex filter is required for the proper frequency separation.

Example 11.2–1 For the circuit shown in Fig. 11.2–4, determine an expression for $v_o(t)$.

Solution. Since the differential pair represents a nonlinear nonmemory device, we may write the Fourier-like series for $i_{C2} = \alpha i_{E2}$ directly from Eqs. (4.6–9) and (4.6–10) in the form

$$i_{C2}(t) \approx i_{E2}(t) = (1\text{ mA})\left\{\frac{1}{2} - \sum_{n=1}^{\infty} a_{2n-1}(x)\cos\left[(2n-1)\left(10^7 t + 10^5\int^t f(\theta)\,d\theta\right)\right]\right\},$$

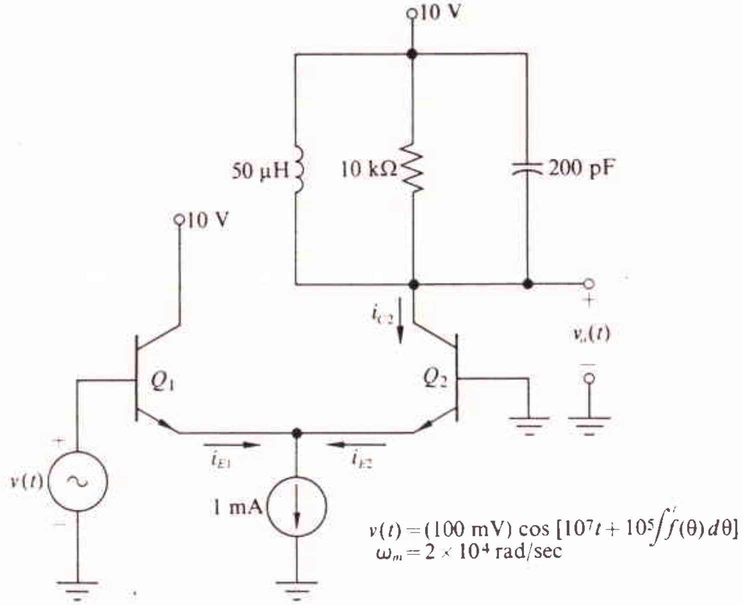

10 V
50 μH
10 kΩ
200 pF
10 V
i_{C2}
+
$v_o(t)$
−
Q_1
Q_2
+
$v(t)$
−
i_{E1}
i_{E2}
1 mA
$v(t) = (100\text{ mV})\cos[10^7 t + 10^5\int^t f(\theta)\,d\theta]$
$\omega_m = 2 \times 10^4$ rad/sec

Figure 11.2–4

where $x = qV_1/kT = 3.85$. The current $i_{C2}(t)$ is clearly the expected superposition of an infinite number of FM signals.

Since $\Delta\omega = 10^5$ rad/sec and $\omega_m = 2 \times 10^4$ rad/sec, $\beta = 5$; therefore, the fundamental FM signal has a bandwidth of $\text{BW}_1 = 3.2 \times 10^5$ rad/sec (cf. Fig. 11.1–4), while the third-harmonic FM signal has a bandwidth of $\text{BW}_3 = 7.5 \times 10^5$ rad/sec. In this ideal case the second-harmonic FM signal is zero. The output filter however, has a center frequency of $\omega_0 = 10^7$ rad/sec, $Q_T = 20$, and a -3 dB bandwidth of 5×10^5 rad/sec. Therefore, the filter passes the fundamental FM signal with only a relatively mild "distortion" while effectively eliminating the harmonic FM signals so that

$$v_o(t) = (10\text{ V}) + (10\text{ V})a_1(x)\cos\left[10^7 t + 10^5\int^t f(\theta)\,d\theta\right].$$

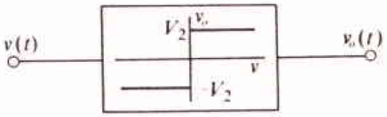
From Fig. 4.6–4 we observe that $a_1(x)$ with $x = 3.85$ has a value of 0.55; thus

$$v_o(t) = (10\text{ V}) + (5.5\text{ V})\cos\left[10^7 t + 10^5\int^t f(\theta)\,d\theta\right].$$

As the amplitude of the FM signal input to the differential pair is increased, we observe from Fig. 4.6–4 that the FM signal in the output remains reasonably constant; thus a bipolar differential pair may be employed as an amplitude limiter when driven by a sufficiently large (with respect to 26 mV) input signal.

onsinusoidal FM Signals

/e have observed that, when an FM signal is passed through a nonlinear device, ie resultant signal contains the same instantaneous frequency information as the riginal FM signal. Therefore, such a distorted FM signal is indeed also an FM gnal, since a signal of the form of Eq. (11.2–1) can be extracted from it by placing it irough either a bandpass filter or an inverse nonlinear operation.



$v(t) = V_1 \cos[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta]$

$v_o(t)$ is an FM square wave

Fig. 11.2–5 Generation of an FM square wave.

For example, if an FM signal is passed through the hard limiter shown in Fig. 1.2–5, an FM square wave of peak amplitude V_2 results. In a similar fashion, if (t) given by Eq. (11.2–1) is passed through a nonlinear device with the transfer haracteristic given by

$$v_o = \frac{2V_2}{\pi} \sin^{-1} \frac{v(t)}{V_1}, \qquad -V_1 \le v(t) \le V_1, \tag{11.2–8}$$

n FM triangular wave of peak amplitude V_2 results. Other nonlinearities yield ther waveforms for the FM signals.

Note, however, that all the nonsinusoidal FM signals, as well as the sinusoidal M signal, have one property in common: when they are plotted vs.

$$\tau(t) = t + \frac{\Delta\omega}{\omega_0} \int^t f(\theta)\, d\theta,$$

hey are periodic with period $2\pi/\omega_0$ [cf. the paragraph containing Eq. (11.2–5)].

Figure 11.2–6 illustrates an FM triangular wave plotted vs. $\tau(t)$ and vs. t. Note hat only when $v_T(t)$ is plotted vs. τ does the waveform consist of a series of straight ine segments.

The fact that a sinusoidal FM signal can be reclaimed from nonsinusoidal FM signals suggests a widely used method of FM generation. Both FM square waves and FM triangular waves are readily generated by standard digital operations. Therefore, FM forms of these waveforms are generated and then converted into a sinusoidal FM signal by filtering (for the FM square wave) or by non-linear processing (for the FM triangular wave). For example, if the FM triangular wave is placed through a nonlinear device with the characteristic

$$v_o = V_1 \sin \frac{\pi v_T(t)}{2V_2},$$

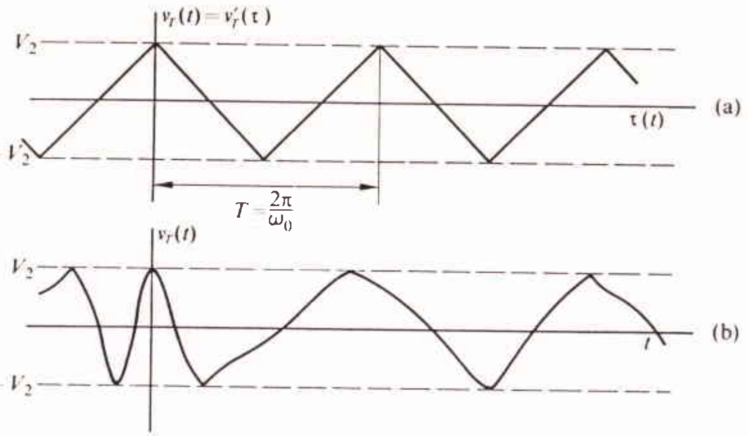


Fig. 11.2–6 FM triangular wave plotted (a) vs. τ and (b) vs. t.

a sinusoidal FM signal results. Whereas filtering to reclaim the sinusoidal FM signal places a limit on the modulation index and deviation ratio that can be used without causing undue distortion, the use of nonlinear processing to reclaim the sinusoidal FM signal imposes no such restrictions.

11.3 TRANSMISSION OF FM SIGNALS THROUGH LINEAR FILTERS

In this section we shall consider the problem of determining the output of a linear filter which is driven by an FM signal of the form

$$v(t) = A \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta\right]. \tag{11.3–1}$$

Unfortunately the solution to such a problem, unlike the corresponding problem in AM, which was studied in Chapter 3, is not possible in closed form for an arbitrary $f(t)$ or an arbitrary filter transfer function $H(p)$. However, two particular cases with wide application in the design of FM systems are tractable mathematically, and we shall restrict our attention to them.

The first tractable situation is that in which $f(t)$ varies slowly in comparison with the duration of the impulse response of the linear filter. In this "quasi-static" case the filter output is obtained by substituting $\omega_i(t) = \omega_0 + \Delta\omega f(t)$ for ω in the filter transfer function $H(j\omega)$ to obtain the output $v_o(t)$ given by

$$v_o(t) = A|H[j\omega_i(t)]| \cos\left\{\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \arg H[j\omega_i(t)]\right\}. \tag{11.3–2}$$

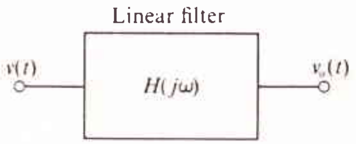
In addition to establishing the validity of this intuitively satisfying "quasi-static" response, we shall attempt to obtain a bound on the maximum value of ω_m [the frequency to which $f(t)$ is band-limited] for which Eq. (11.3–2) may be employed.

The second tractable situation is that in which $H(j\omega)$ has both a magnitude and a phase which are straight line functions of ω. In this case, regardless of the rate at which $f(t)$ is varied, $v_o(t)$ is given by

$$v_o(t) = |H[j\omega_i(t - t_0)]| \cos\left[\omega_0 t + \Delta\omega \int^{t-t_0} f(\theta)\,d\theta + \theta(\omega_0)\right], \qquad (11.3\text{–}3)$$

where $\theta(\omega_0) = \arg H(j\omega_0)$ and t_0 is the constant slope of the arg $H(j\omega)$-vs.-ω curve. We observe from Eq. (11.3–3) that $|H(j\omega)|$ establishes the transfer function between the envelope of $v_o(t)$ and $\omega_i(t - t_0)$, whereas in the "quasi-static" case $|H(j\omega)|$ establishes the transfer function between the envelope of $v_o(t)$ and $\omega_i(t)$. Although Eqs. (11.3–2) and (11.3–3) seem to be quite different in form, it is readily shown that Eq. (11.3–3) reduces to Eq. (11.3–2) as ω_m decreases relative to $1/t_0$.

In either of the above cases it is clear that, if $|H(j\omega)|$ is a straight line function of ω, the filter output $v_o(t)$ has an envelope which varies in proportion to $\omega_i(t)$; thus if we place an envelope detector at the filter output we obtain an obvious means of FM demodulation. This technique will be discussed in more detail in Chapter 12.



$h(t) = \mathfrak{F}^{-1}H(j\omega)$

$v(t) = A\mathcal{R}e\,(\exp j\omega_0 t)[\exp j\Delta\omega \int^t f(\theta)\,d\theta]$

Fig. 11.3–1 Linear filter driven by FM signal.

To demonstrate the validity of Eq. (11.3–2) consider the filter driven by $v(t)$ which is shown in Fig. 11.3–1. If $h(t)$ is the impulse response of the filter and we represent the filter input $v(t)$ in the form

$$v(t) = A\cos[\omega_0 t + \psi(t)] = A\mathcal{R}e\, e^{j\omega_0 t} e^{j\psi(t)}, \qquad (11.3\text{–}4)$$

where $\mathcal{R}e$ denotes the real part and $\psi(t) = \Delta\omega \int^t f(\theta)\,d\theta$, then employing the convolution integral we may write $v_o(t)$ as

$$v_o(t) = A\,\mathcal{R}e\left[e^{j\omega_0 t}\int_{-\infty}^{\infty} h(\tau)e^{-j\omega_0\tau}e^{j\psi(t-\tau)}\,d\tau\right]. \qquad (11.3\text{–}5)$$

If we now expand $\psi(t - \tau)$ in a MacLauren series in τ, we obtain

$$\psi(t-\tau) = \psi(t) - \dot{\psi}(t)\tau + \frac{\ddot{\psi}(t)\tau^2}{2!} - \cdots, \qquad (11.3\text{–}6)$$

where the dots represent differentiation with respect to t (or $-\tau$). However, if $h(\tau)$ is zero for $\tau < 0$ and decays to zero for $\tau > \tau_M$ (for example, τ_M might represent four time constants if $h(t)$ were the impulse response of a single-pole filter), then $\psi(t - \tau)$ contributes to the integral of Eq. (11.3–5) only for $0 < \tau < \tau_M$. Thus if

$$\frac{\ddot{\psi}(t)\tau_M^2}{2} \ll \dot{\psi}\tau_M$$

(that is, if the rate of change of $\dot{\psi}$, the ac portion of the instantaneous frequency, is small, i.e., if $\omega_m \ll 1/\tau_M$), then Eq. (11.3–5) simplifies to

$$\begin{aligned} v_o(t) &= A\mathcal{R}e\left\{e^{j[\omega_0 t + \psi(t)]}\int_{-\infty}^{\infty} h(\tau)e^{-j\overbrace{[\omega_0 + \dot{\psi}(t)]}^{\omega_i(t)}\tau}\,d\tau\right\} \\ &= A\mathcal{R}e\{e^{j[\omega_0 t + \psi(t)]}H[j\omega_i(t)]\} \\ &= A|H[j\omega_i(t)]|\cos\{\omega_0 t + \psi(t) + \arg H[j\omega_i(t)]\}, \end{aligned} \qquad (11.3\text{–}7)$$

which is the desired result. In obtaining Eq. (11.3–7) use was made of the fact that the defining equation for $H(j\omega)$ is the Fourier transform of the impulse response given by

$$H(j\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau}\,d\tau. \qquad (11.3\text{–}8)$$

To obtain a limit on the usefulness of Eq. (11.3–7) we shall follow an approach suggested by Baghdady.† We first employ the expansion of Eq. (11.3–6) to write

$$\exp j\psi(t-\tau) = [\exp j\psi(t)][\exp - j\dot{\psi}(t)\tau]\exp j\left[\frac{\ddot{\psi}(t)\tau^2}{2!} - \frac{\dddot{\psi}(t)\tau^3}{3!} + \cdots\right]. \qquad (11.3\text{–}9)$$

We then further expand the last term of Eq. (11.3–9) in a MacLauren series in τ to obtain

$$\exp j\psi(t-\tau) = [\exp j\psi(t)][\exp - j\dot{\psi}(t)\tau]\left[1 + j\frac{\ddot{\psi}(t)\tau^2}{2!} - j\frac{\dddot{\psi}(t)\tau^3}{3!} + \cdots\right]. \qquad (11.3\text{–}10)$$

Finally, if we assume that for small values of τ_M the $j[\ddot{\psi}(t)\tau^2/2!]$-term is the main term contributing to inaccuracies in Eq. (11.3–7) (which is usually the case in practice, even though some nonrealistic examples can be set up where $\dddot{\psi}(t)\tau^3/3!$ is more significant than $\ddot{\psi}(t)\tau^2/2!$), then Eq. (11.3–9) may be written in the form of the desired "quasi-static" output plus an error term:

$$\begin{aligned} v_o(t) = A\mathcal{R}e\Bigg\{e^{j[\omega_0 t + \psi(t)]}\Bigg[&\int_{-\infty}^{\infty} h(\tau)e^{-j[\omega_0 + \dot{\psi}(t)]\tau}\,d\tau \\ &+ j\frac{\ddot{\psi}(t)}{2}\int_{-\infty}^{\infty}\tau^2 h(\tau)e^{-j[\omega_0 + \dot{\psi}(t)]\tau}\,d\tau\Bigg]\Bigg\}. \end{aligned} \qquad (11.3\text{–}11)$$

---

† E. J. Baghdady, *Lectures on Communication System Theory*, McGraw-Hill, New York (1961), pp. 472–483.

Since multiplication in the t- (or τ-) domain corresponds to differentiation in the frequency domain

$$\tau^2 h(\tau) \leftrightarrow -\frac{d^2H(j\omega)}{d\omega^2}. \tag{11.3–12}$$

Equation (11.3–11) may therefore be rewritten in the form

$$v_o(t) = A\mathcal{R}e\left\{e^{j[\omega_0 t+\psi(t)]}H[j\omega_i(t)]\left[1 - j\frac{\ddot{\psi}(t)}{2}\frac{\dfrac{d^2H[j\omega_i(t)]}{d\omega_i^2}}{H[j\omega_i(t)]}\right]\right\}. \tag{11.3–13}$$

It is now apparent that if

$$\left|\frac{\ddot{\psi}(t)}{2}\frac{\dfrac{d^2H[j\omega_i(t)]}{d\omega_i^2}}{H[j\omega_i(t)]}\right|_{\max} \le \left|\frac{\ddot{\psi}(t)}{2}\right|_{\max}\left|\frac{\dfrac{d^2H[j\omega_i(t)]}{d\omega_i^2}}{H[j\omega_i(t)]}\right|_{\max} \ll 1, \tag{11.3–14}$$

then Eq. (11.3–13) reduces to Eq. (11.3–7) and the "quasi-static" approximation is valid.

As an example of applying Eq. (11.3–4), consider passing an FM signal through a high-Q_T parallel RLC circuit for which $Z_{11}(j\omega)$ is closely approximated by (cf. Eq. 2.2–9)

$$Z_{11}(j\omega) = \frac{R}{1 + j\dfrac{\omega-\omega_0}{\alpha}}, \qquad \omega > 0,$$

where $\omega_0 = 1/\sqrt{LC}$ and $\alpha = 1/2RC$ (the reciprocal of the circuit time constant). It follows immediately that

$$\frac{d^2Z_{11}(j\omega)}{d\omega^2} = \frac{-2R}{\alpha^2\left(1 + j\dfrac{\omega-\omega_0}{\alpha}\right)^3} \tag{11.3–15}$$

and

$$\left|\frac{\dfrac{d^2Z_{11}(j\omega)}{d\omega^2}}{Z_{11}(j\omega)}\right|_{\max} = \frac{2}{\alpha^2}. \tag{11.3–16}$$

In deriving Eq. (11.3–16) we have assumed that $\omega_i(t)$ is centered about ω_0. If this were not the case, Eq. (11.3–16) would give a smaller value.

If we now let $f(t) = \cos\omega_m t$ (the worst case), then

$$\dot{\psi}(t) = \Delta\omega\cos\omega_m t, \qquad \ddot{\psi}(t) = -\Delta\omega\omega_m\sin\omega_m t,$$

and

$$|\ddot{\psi}(t)|_{\max} = \Delta\omega\omega_m;$$

therefore, the quasi-static approximation is valid provided that

$$\frac{\Delta\omega\omega_m}{\alpha^2} = \frac{1}{\beta}\left(\frac{\Delta\omega}{\alpha}\right)^2 \ll 1. \tag{11.3–17}$$

If we interpret "much less than unity" as 1/20 (5% error or less), then for $\beta = \Delta\omega/\omega_m = 5$,

$$\frac{\Delta\omega}{\alpha} \le \frac{1}{2},$$

and for $\beta = 10$,

$$\frac{\Delta\omega}{\alpha} \le \frac{1}{\sqrt{2}}.$$

However, since the derivation for the maximum value of $\Delta\omega/\alpha$ involved consistently conservative approximations, in most cases the bounds may be relaxed by 50% and the quasi-static approach will still yield reasonably accurate results.

Note from Eq. (11.3–17) that the "quasi-static" approximation is valid if either $\omega_m \ll \alpha$ or $\Delta\omega \ll \alpha$. Therefore, if the maximum modulation rate of $f(t)$ [or $\omega_i(t)$] is much less than the reciprocal of the circuit time constant, the deviation may be large without destroying the validity of Eq. (11.3–7). On the other hand, if $2\Delta\omega \ll 2\alpha$, where 2α is the circuit 3 dB bandwidth, then $f(t)$ may vary quite rapidly without

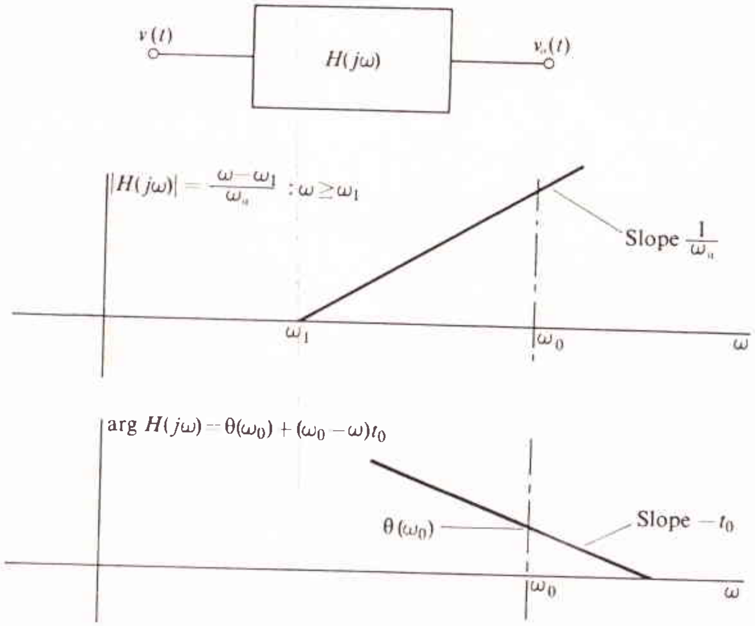

$v(t)$
$H(j\omega)$
$v_o(t)$
$|H(j\omega)| = \frac{\omega-\omega_1}{\omega_a}$; $\omega \ge \omega_1$
Slope $\frac{1}{\omega_a}$
ω_1
ω_0
ω
$\arg H(j\omega) = \theta(\omega_0) + (\omega_0 - \omega)t_0$
$\theta(\omega_0)$
Slope $-t_0$
ω_0
ω

Fig. 11.3–2 Filter with linearly sloping magnitude and phase characteristic.

destroying the validity of the "quasi-static" solution. Although we derived Eq. (11.3–17) for a specific filter, a similar equation results for almost all narrowband filters; specifically,

$$\frac{K}{\beta}\left(\frac{2\,\Delta\omega}{\mathrm{BW}}\right)^2 \ll 1, \tag{11.3–18}$$

where K usually lies between 1 and 5. Several more specific examples are derived by Baghdady.[†]

To demonstrate the validity of Eq. (11.3–3), let us consider the filter with the magnitude and phase function shown in Fig. 11.3–2. If $v(t)$ is again given by

$$v(t) = \mathscr{R}e[Ae^{j[\omega_0 t + \psi(t)]}],$$

then

$$v_o(t) = \mathscr{R}e[\mathscr{F}^{-1}H(j\omega)\tilde{V}(\omega)], \tag{11.3–19}$$

where $\tilde{V}(\omega)$ is the Fourier transform of $Ae^{j[\omega_0 t + \psi(t)]}$. With $H(j\omega)$ given in Fig. 11.3–2 substituted into Eq. (11.3–19), $v_o(t)$ reduces to

$$\begin{aligned} v_o(t) &= \mathscr{R}e\left\{\frac{e^{j\theta(\omega_0)}e^{j\omega_0 t_0}}{\omega_z}\mathscr{F}^{-1}\left[\frac{j\omega\tilde{V}(\omega)e^{-j\omega t_0}}{j} - \omega_1\tilde{V}(\omega)e^{-j\omega t_0}\right]\right\} \\ &= \mathscr{R}e\left\{\frac{Ae^{j\theta(\omega_0)}e^{j\omega_0 t_0}}{\omega_z}\left[\frac{de^{j[\omega_0(t-t_0)+\psi(t-t_0)]}}{j\,dt} - \omega_1 e^{j[\omega(t-t_0)+\psi(t-t_0)]}\right]\right\} \\ &= \mathscr{R}e\,A\frac{\overbrace{\omega_0 + \dot{\psi}(t-t_0)}^{\omega_i(t-t_0)} - \omega_1}{\omega_z}e^{j[\omega_0 t + \psi(t-t_0) + \theta(\omega_0)]} \\ &= A|H[j\omega_i(t-t_0)]|\cos[\omega_0 t + \psi(t-t_0) + \theta(\omega_0)], \end{aligned} \tag{11.3–20}$$

which is the desired result. In the above derivation, use has been made of the fact that the operations of $e^{-j\omega t_0}$ and $j\omega$ in the frequency domain are equivalent to a time delay of t_0 and the differentiation operation respectively in the time domain.

11.4 FREQUENCY MODULATION TECHNIQUES—THE FM DIFFERENTIAL EQUATION

In this section and the sections to follow we shall study the theoretical methods (and their practical implementation) by which we may vary the instantaneous frequency of a sinusoidal carrier in proportion to the modulation information $f(t)$ to obtain an FM signal of the form

$$v(t) = A\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right] \tag{11.4–1}$$

which has the instantaneous frequency

$$\omega_i(t) = \omega_0 + \Delta\omega f(t). \tag{11.4–2}$$

[†] *Op. cit.*

In general, the theoretical methods of producing a frequency-modulated signal can be broken into six basic categories:

1. The analog simulation of the FM differential equation.
2. The quasi-static variation of the oscillation frequency of an oscillator.
3. The generation and nonlinear waveshaping of an FM triangular wave.
4. The generation and filtering of an FM square wave.
5. The generation and frequency multiplication of narrowband FM (the Armstrong method).
6. The generation of FM in specialized devices, e.g. the Phasitron etc.

Of all the above methods only the analog simulation of the FM differential equation and the generation and waveshaping of a triangular FM wave produce a theoretically perfect FM signal which may have a frequency deviation $\Delta\omega$ or maximum modulation frequency ω_m as high as the carrier frequency ω_0. Both the generation and filtering of an FM square wave and the Armstrong method require the filtering of a distorted FM signal to produce the output FM signal and thus have $\Delta\omega$ and ω_m limited with respect to ω_0 by curves similar to those of Fig. 11.2–3. In the quasi-static method both $\Delta\omega$ and ω_m are even more severely restricted. The maximum modulation frequency must be kept small in comparison with the reciprocal of the envelope transient time constant of the oscillator, while $\Delta\omega$ must be much less than ω_0 to prevent nonlinearities in the instantaneous frequency of the FM signal.

In the remainder of this section and the sections following we shall consider each of the above techniques in somewhat greater detail. Specifically, we shall attempt to obtain a block diagram for an FM generator corresponding to each of the above methods, along with the maximum values of $\Delta\omega$ and ω_m (relative to ω_0) for which the generator functions properly. In each case we shall then develop one or more physical implementations for the generator.

The FM Differential Equation[†]

A second-order, linear, homogeneous differential equation which has as its two linearly independent solutions the FM signals

$$v_1 = \cos\int^t \omega_i(\theta)\,d\theta \qquad \text{and} \qquad v_2 = \sin\int^t \omega_i(\theta)\,d\theta \tag{11.4–3}$$

is given in integro-differential form by

$$\omega_i(t)\int^t \omega_i(\theta)v(\theta)\,d\theta + \dot{v}(t) = 0 \tag{11.4–4}$$

or, equivalently, is given in differential form by

$$v(t) - \frac{\dot{v}(t)\dot{\omega}_i(t)}{\omega_i(t)^3} + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0. \tag{11.4–5}$$

[†] D. T. Hess, "FM Differential Equation" (letter), *Proc. IEEE*, **54**, No. 8, p. 1089 (Aug. 1966).

Direct substitution of v_1 and v_2 given by Eq. (11.4–3) into Eq. (11.4–4) or Eq. (11.4–5) indicates that v_1 and v_2 are indeed two independent solutions; and since the differential equation is second-order and linear [as clearly indicated by Eq. (11.4–5)], any other solution may be obtained as a linear combination of v_1 and v_2. Hence the most general solution may be written as

$$v(t) = A_1 \cos \int^t \omega_i(\theta)\, d\theta + A_2 \sin \int^t \omega_i(\theta)\, d\theta$$

$$= C \cos \left[\int^t \omega_i(\theta)\, d\theta + \theta_0 \right], \tag{11.4–6}$$

where A_1, A_2, C, and θ_0 are arbitrary constants of integration which may be uniquely specified by the initial conditions $v(0)$ and $\dot{v}(0)$.

For example if $v(0) = V_1$ and $\dot{v}(0) = 0$, then

$$v(t) = V_1 \cos \int^t \omega_i(\theta)\, d\theta;$$

and if $\omega_i(t) = \omega_0 + \Delta\omega f(t)$, then

$$v(t) = V_1 \cos \left[\omega_0 t + \Delta\omega \int_0^t f(\theta)\, d\theta \right], \tag{11.4–7}$$

which is the desired FM signal. Any other set of initial conditions merely changes the amplitude of $v(t)$ and introduces a constant phase angle which is of little consequence, since it may be taken as zero with a different choice of time axis.

Note that Eq. (11.4–4) and (11.4–5) are even functions in $\omega_i(t)$; hence negative values of $\omega_i(t)$ produce the same $v(t)$ as the corresponding positive values of $\omega_i(t)$. Thus to avoid this full-wave rectification effect of $\omega_i(t)$, the instantaneous frequency must be greater than zero; that is,

$$\omega_i(t) = \omega_0 + \Delta\omega f(t) > 0 \tag{11.4–8}$$

or equivalently

$$\Delta\omega < \omega_0. \tag{11.4–9}$$

This restriction is of little consequence, since no demodulator exists which can recover $f(t)$ if Eq. (11.4–9) is not satisfied.

Generation of an FM signal by simulating the FM differential equation is particularly well suited to an analog computer. Figure 11.4–1 illustrates an analog computer block diagram which employs two multipliers, two integrators, and an inverter to implement Eq. (11.4–4). For the block diagram of Fig. 11.4–1 the equation describing $v(t)$ takes the form

$$K_M K_I v_i(t) \int^t K_M K_I v_i(\theta) v(\theta)\, d\theta + \dot{v}(t) = 0, \tag{11.4–10}$$

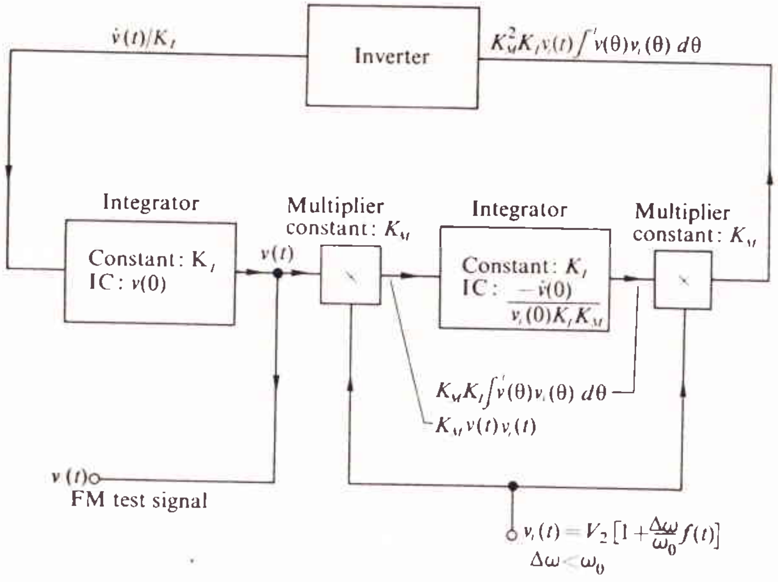

$\dot{v}(t)/K_I$
Inverter
$K_M^2 K_I v_i(t) \int^t v(\theta) v_i(\theta)\, d\theta$
Integrator
Constant: K_I
IC: $v(0)$
$v(t)$
Multiplier
constant: K_M
Integrator
Constant: K_I
IC: $\frac{-\dot{v}(0)}{v_i(0) K_I K_M}$
Multiplier
constant: K_M
$K_M K_I \int^t v(\theta) v_i(\theta)\, d\theta$
$K_M v(t) v_i(t)$
$v(t)$
FM test signal
$v_i(t) = V_2 [1 + \frac{\Delta\omega}{\omega_0} f(t)]$
$\Delta\omega < \omega_0$

Fig. 11.4–1 Analog computer implementation of FM differential equation.

which is identical in form to Eq. (11.4–4). It is thus apparent that the instantaneous frequency of $v(t)$ in the block diagram of Fig. 11.4–1 is

$$\omega_i(t) = K_I K_M v_i(t). \tag{11.4–11}$$

Hence, if $v_i(t)$ is chosen as

$$V_2 \left[1 + \frac{\Delta\omega}{\omega_0} f(t) \right]$$

and $K_I K_M V_2$ is adjusted to equal ω_0, the analog computer has the desired instantaneous frequency $\omega_i(t) = \omega_0 + \Delta\omega f(t)$.

If, in addition the initial conditions (IC) of the integrators are set so that

$$v(0) = V_1 \qquad \text{and} \qquad \frac{-\dot{v}(0)}{K_I K_M v_i(0)} = 0,$$

with the result that $\dot{v}(0) = 0$, then the analog computer generates the FM signal given by Eq. (11.4–7). (A practical circuit to take care of the initial-condition problem automatically will be presented shortly.)

The block diagram of Fig. 11.4–1 can be constructed not only on the analog computer, but also as an independent FM generator. If good-quality integrated operational amplifiers are employed as the basic component of the integrators and the inverter and if high-speed integrated multipliers (Section 8.3) are employed, then the resultant FM generator should perform satisfactorily with carrier frequencies exceeding 10 MHz. In addition, linear frequency deviations equal to the carrier frequency may be obtained regardless of the modulation frequency, which may exceed the carrier frequency. The extreme range of $\Delta\omega$ and ω_m is a direct result of the fact that the solution of the FM differential equation is a theoretically perfect FM signal regardless of the values of $\Delta\omega$ and ω_m, provided that $\Delta\omega < \omega_0$.

Because of the high deviation ratio ($D = \Delta\omega/\omega_0$) possible with the analog modulator of Fig. 11.4–2, an FM signal with specified values of $\Delta\omega$ and ω_m can be generated at a reasonably low carrier frequency ω_0. This FM signal can then be heterodyned by a crystal oscillator with frequency ω_c to any desired carrier frequency $\omega_0' = \omega_0 + \omega_c$. If $\omega_c \gg \omega_0$, as is the case when ω_0 was initially a low value, then the stability of the center frequency of the resultant FM signal is controlled primarily by the stable crystal oscillator and is almost independent of the less stable frequency ω_0. For example, if $\omega_c = 9\omega_0$ and ω_0 varies by 1%, then ω_c varies by only 0.1%.

The limit on the maximum ratio of ω_c to ω_0 to obtain center frequency stability is set by the problem of separating the FM signal at the difference frequency $\omega_c - \omega_0$ (heterodyning produces both the sum and the difference frequencies) from the desired FM signal. Specifically, a filter with BW $\ll 2\omega_0$ and a center frequency of $\omega_c + \omega_0$ is required to perform the separation. The Q_T of such a filter must satisfy the relationship

$$Q_T = \frac{\omega_0'}{\text{BW}} \gg \frac{1}{2}\left(\frac{\omega_c}{\omega_0} + 1\right); \qquad (11.4\text{–}12)$$

hence, as ω_c/ω_0 increases, the required Q_T and thus the filter complexity greatly increase.

Amplitude Limiting and Initial Conditions

To avoid the problem of having to set the initial conditions to ensure an FM signal of a desired amplitude and, in turn, to extend the usefulness of the FM generator beyond the low-frequency analog computer two additional feedback loops may be added to the block diagram of Fig. 11.4–1 as shown in Fig. 11.4–2. The two additional loops have the same effect on the FM generator as an amplitude-limiting circuit has on a sinusoidal oscillator. As a matter of fact, with $v_i(t) = V_2$ (a constant), the effects are identical. Specifically, for values of $V_{\text{env}} < V_1$, the system poles lies in the right-half complex plane and a growing oscillation occurs. When $V_{\text{env}} = V_1$, the poles lie on the imaginary axis and the oscillations stabilize. If, however, for some reason V_{env} increases beyond V_1, the complex conjugate poles move into the left-half plane and the envelope of the oscillation decays back toward V_1.

It is apparent that even if $v_i(t)$ is a function of time, the envelope of $v(t)$ should stabilize at V_1, since for this value no signal is fed into multipliers M_3 and M_4, the two additional loops are opened, and the block diagram of Fig. 11.4–2 reduces to the original diagram of Fig. 11.4–1.

To obtain a more quantitative picture of how the envelope of $v(t)$ stabilizes, we may write the differential equation governing the behavior of $v(t)$ from the block

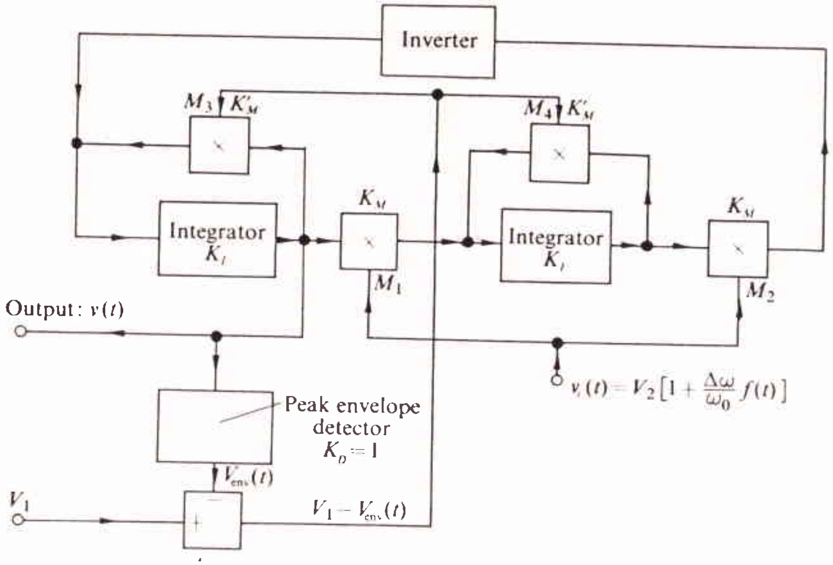


Fig. 11.4–2 Analog implementation of FM differential equation with amplitude stabilization mechanism.

diagram of Fig. 11.4–2†:

$$0 = v(t)\left(1 + \frac{\alpha\dot{\omega}_i}{\omega_i^3} + \frac{\alpha^2}{\omega_i^2} - \frac{\dot{\alpha}}{\omega_i^2}\right) - \dot{v}(t)\left(\frac{2\alpha}{\omega_i^2} + \frac{\dot{\omega}_i}{\omega_i^3}\right) + \frac{\ddot{v}(t)}{\omega_i^2}, \qquad (11.4\text{–}13)$$

where

$$\omega_i(t) = K_I K_M v_i(t) \qquad \text{and} \qquad \alpha(t) = K_I K_M'[V_1 - V_{\text{env}}(t)].$$

For this case, the most general solution is given [as the reader should verify by substituting in Eq. (11.4–13)] by

$$v(t) = C\underbrace{\left[\exp \int^t \alpha(\theta)\, d\theta\right]}_{V_{\text{env}}(t)} \cos\left[\int^t \omega_i(\theta)\, d\theta + \theta_0\right], \qquad (11.4\text{–}14)$$

where again C and θ_0 are arbitrary constants.

It is apparent that from Eq. (11.4–14) that

$$V_{\text{env}}(t) = C \exp \int^t K_I K_M'[V_1 - V_{\text{env}}(\theta)]\, d\theta, \qquad (11.4\text{–}15)$$

---

† Equation (11.4–13) is not obvious by inspection. The interested reader should make an effort to derive it from the block diagram of Fig. 11.4–2; the more casual reader should accept its validity and proceed to the conclusion drawn from the equation.

or equivalently,

$$\ln \frac{V_{\text{env}}(t)}{C} = \int^{t} K_I K_M'[V_1 - V_{\text{env}}(\theta)]\, d\theta. \tag{11.4–16}$$

By differentiating Eq. (11.4–16) with respect to time, we obtain

$$\frac{dV_{\text{env}}(t)/dt}{V_{\text{env}}(t)[V_1 - V_{\text{env}}(t)]} = K_I K_M', \tag{11.4–17}$$

which has as its solution

$$V_{\text{env}}(t) = \frac{V_1}{1 + Be^{-K_M' V_1 K_I t}}, \tag{11.4–18}$$

where B is an arbitrary constant of integration. As expected, we observe from Eq. (11.4–17) that $V_{\text{env}}(t)$ approaches V_1 and does so with a time constant $\tau = 1/K_M' V_1 K_I$. Therefore, if power is applied at $t = 0$, for $t > 4\tau$, $v(t)$ (for Fig. 11.4–2) is given by

$$v(t) = V_1 \cos\left[\omega_0 t + \Delta\omega \int^{t} f(\theta)\, d\theta\right].$$

In the above analysis we assumed that the envelope detector extracted the instantaneous envelope of $v(t)$. In general, however, the detector contains some filtering, so that the detector output is a filtered version of the instantaneous envelope. If this equivalent filter contains a single pole at some low frequency (which most peak envelope detectors do), then the above results are still valid except that $V_{\text{env}}(t)$ rises toward V_1 more slowly and with a second-order response which is usually dominated by the low-frequency detector pole. If, on the other hand, the peak envelope detector contains two or more poles, the possibility of squegging exists; hence higher-order envelope detector filters should be avoided.

In addition, the single pole of the peak envelope detector should lie well below the lowest value of the instantaneous frequency of prevent the possibility of having the envelope detector output vary over a carrier cycle. Such a variation would significantly distort the FM signal, since $\alpha(t)$ in Eq. (11.4–13) would vary over a cycle in this case.

11.5 QUASI-STATIC FREQUENCY MODULATION

When the frequency deviation $\Delta\omega$ and the maximum modulation frequency ω_m of an FM signal are both small fractions of the carrier frequency ω_0 (which is the case in more than 95% of the practical applications for which an FM signal is required), then the FM differential equation given by Eq. (11.4–5) can be closely approximated by either of the following quasi-static forms, which act as the defining equations for a broad class of FM generators:

$$v(t) + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0 \tag{11.5–1a}$$

and

$$v(t) - \frac{2\dot{v}(t)\dot{\omega}_i(t)}{\omega_i^3} + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0. \tag{11.5–1b}$$

The justification for this approximation is based on the fact that, when both $\Delta\omega \ll \omega_0$ and $\omega_m \ll \omega_0$, the term $-\dot{v}(t)\dot{\omega}_i(t)/\omega_i(t)^3$ in Eq. (11.4–5) becomes vanishingly small in comparison with the terms $v(t)$ and $\ddot{v}(t)/\omega_i(t)^2$; hence it may be added to (Eq. 11.5–1b) or subtracted from (Eq. 11.5–1a) Eq. (11.4–5) with negligible effect on the solution $v(t)$.†

For example, if we consider the case where $\omega_i(t) = \omega_0 + \Delta\omega \cos \omega_m t$ (the most restrictive case for given values $\Delta\omega$, ω_m, and ω_0), then the solution to the FM differential equation is given by

$$v(t) = A \cos\left(\omega_0 t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t + \theta_0\right),$$

from which it follows that

$$\dot{v}(t) = -A(\omega_0 + \Delta\omega \cos \omega_m t) \sin\left(\omega_0 t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t + \theta_0\right) \tag{11.5–2}$$

where A and θ_0 are arbitrary constants of integration. Therefore, from Eq. (11.5–2) and the fact that $\dot{\omega}_i(t) = -\Delta\omega\omega_m \sin \omega_m t$, it is apparent that

$$\frac{\dot{v}\dot{\omega}_i}{\omega_i^3} = \frac{A\Delta\omega\omega_m \sin \omega_m t}{(\omega_0 + \Delta\omega \cos \omega_m t)^2} \sin\left(\omega_0 t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t + \theta_0\right) \tag{11.5–3}$$

and thus

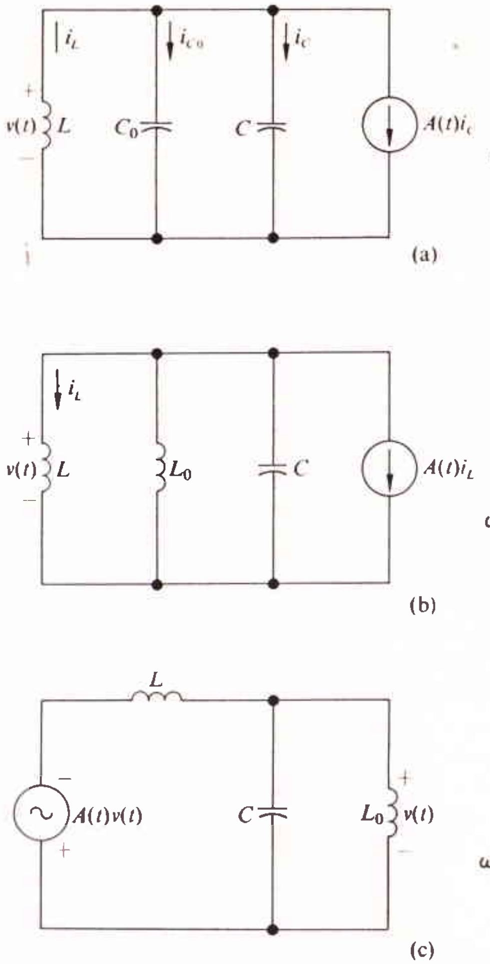
$$\frac{|\dot{v}\dot{\omega}_i/\omega_i^3|_{\max}}{|v|_{\max}} \approx \frac{\Delta\omega\omega_m}{\omega_0^2} = \frac{1}{\beta}\left(\frac{\Delta\omega}{\omega_0}\right)^2 \tag{11.5–4}$$

for $\Delta\omega/\omega_0 \ll 1$. For $\beta = \Delta\omega/\omega_m = 5$ and $\omega_0 = 20\,\Delta\omega$, Eq. (11.5–4) has the numerical value of 5×10^{-4}, for which the second term in the FM differential equation relative to the first term is truly negligible. In most practical modulators, as we shall see, ω_0 is usually much greater than $20\,\Delta\omega$; thus Eq. (11.5–1a and b) is an excellent approximation to the FM differential equation.

The advantage of being able to employ Eq. (11.5–1a and b) as the basis for producing an FM signal is that it is much more simply implemented by practical circuits than is the exact equation. Specifically, each of the *LC* circuit models shown in Fig. 11.5–1 has as its defining differential equation one of the quasi-static FM differential

---

† If a negligibly small term is added to or subtracted from the FM differential equation having a set of initial conditions specified at $t = 0$, it will be a considerable length of time before the solutions to the original differential equation and the modified differential equations differ significantly. However, if the circuit which implements the modified differential equation contains an envelope-limiting mechanism which is fast-acting in comparison with ω_m (all practical FM generators must contain such a mechanism), then the initial conditions of the modified differential equation are updated at a sufficiently rapid rate to ensure that its solution differs from the solution of the original differential equation by a negligible amount.



$$v(t) - \frac{2\dot{\omega}_i(t)\dot{v}(t)}{\omega_i(t)^3} + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0$$

$$\omega_i(t) = \frac{1}{\sqrt{L(C+C_0)\left(1+\dfrac{CA(t)}{C+C_0}\right)}}$$

$$v(t) - \frac{2\dot{\omega}_i(t)\dot{v}(t)}{\omega_i(t)^3} + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0$$

$$\omega_i(t) = \sqrt{\frac{L_0+L}{CLL_0}\left(1+\frac{L_0A(t)}{L_0+L}\right)}$$

$$v(t) + \frac{\ddot{v}(t)}{\omega_i(t)^2} = 0$$

$$\omega_i(t) = \sqrt{\frac{L_0+L}{CLL_0}\left(1+\frac{L_0A(t)}{L_0+L}\right)}$$

Fig. 11.5–1 Three quasi-static FM modulators.

equations. In each circuit a time-varying controlled source is employed to vary the instantaneous frequency. The imaginative reader should be able to construct several other LC circuits with time-varying controlled sources which also have as their defining equations one of the quasi-static FM equations.

To obtain the differential equation for any of the circuits of Fig. 11.5–1, one must write a single-node equation. For example, for circuit (a) we have

$$i_L + i_{C0} + [1 + A(t)]i_C = 0$$

$$= \frac{1}{L}\int^t v(\theta)\,d\theta + C_0\dot{v}(t) + C[1 + A(t)]\dot{v}(t) \qquad (11.5\text{–}5)$$

or equivalently,

$$0 = \int^t v(\theta)\,d\theta + \frac{\ddot{v}(t)}{\omega_i(t)^2}, \qquad (11.5\text{–}6)$$

where

$$\omega_i(t) = \frac{1}{\sqrt{L(C_0 + C)\left[1 + \dfrac{CA(t)}{C_0 + C}\right]}}.$$

If we differentiate Eq. (11.5–6) with respect to time we obtain Eq. (11.5–1b), which is the desired result.

As an alternative point of view, the time-varying controlled source in the circuits of Fig. 11.5–1 may be thought of as varying the capacitance or inductance of the circuit model as a function of time. For example, the last term in Eq. (11.5–5) could be regarded as the current of a time-varying capacitance $C(t) = [1 + A(t)]C$. With this point of view, we obtain an additional technique for implementing quasi-static FM generators which entails placing time-varying capacitors or inductors in shunt with parallel LC circuits.

In each of the circuits of Fig. 11.5–1 we note that $\omega_i(t)$ is a nonlinear function of $A(t)$; thus if $A(t)$ varies in proportion to the modulation information $f(t)$, that is, if

$$A(t) = A_0 + A_1 f(t), \qquad (11.5\text{–}7)$$

then $A_1 f(t)$ must be kept sufficiently small so that $\omega_i(t)$ also varies in proportion to $f(t)$, as is usually desired. Specifically, if for circuit (a) of Fig. 11.5–1 we expand $\omega_i(t)$ in a MacLauren expansion in $A_1 f(t)$ to obtain

$$\omega_i(t) = \omega_0\left[1 + \frac{\Delta\omega}{\omega_0}f(t) + \frac{3}{4}\left(\frac{\Delta\omega}{\omega_0}\right)^2 f^2(t) + \cdots\right], \qquad (11.5\text{–}8)$$

where

$$\omega_0 = \frac{1}{\sqrt{L(C_0 + C)\left(1 + \dfrac{CA_0}{C_0 + C}\right)}}$$

and

$$\frac{\Delta\omega}{\omega_0} = \frac{-CA_1}{2(C_0 + C)}\,\frac{1}{[1 + CA_0/(C_0 + C)]},$$

then we observe that only if

$$\frac{3}{4}\left(\frac{\Delta\omega}{\omega_0}\right)^2 \ll \frac{\Delta\omega}{\omega_0} \qquad (11.5\text{–}9)$$

does the instantaneous frequency have the desired form $\omega_i(t) = \omega_0 + \Delta\omega f(t)$. For example, if the maximum amplitude of the nonlinear term of Eq. (11.5–8) is to be less than 1% of the maximum amplitude of the linear term, then $\Delta\omega/\omega_0 < 0.0133$ or equivalently $\omega_0 > 75\,\Delta\omega$. If $A_0 \ll 1$ (which is usually true in practice), then

$$\omega_0 \approx \frac{1}{\sqrt{L(C_0 + C)}} \qquad \text{and} \qquad \frac{\Delta\omega}{\omega_0} \approx -\frac{CA_1}{2(C_0 + C)}.$$

Identical restrictions on $\Delta\omega/\omega_0$ are obtained for the other circuits of Fig. 11.5–1; hence, if $\beta = \Delta\omega/\omega_m > 1$ (which is usually true for the majority of FM signals) then satisfying Eq. (11.5–9) is sufficient to ensure that Eq. (11.5–1) is also valid. Therefore, with $\beta > 1$ and $\omega_0/\Delta\omega > 75$, the output signals of the quasi-static FM generators are of the desired form

$$\begin{aligned} v(t) &= A\cos\left[\int^t \omega_i(\theta)\,d\theta + \theta_0\right] \\ &= A\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta + \theta_0\right]. \end{aligned} \qquad (11.5\text{–}10)$$

The problem now remains of synthesizing the circuit models of Fig. 11.5–1 with practical circuits. The first obstacle to be overcome is the synthesis of a lossless LC circuit. This is readily accomplished by placing a lossy LC circuit which has the form of the circuits of Fig. 11.5–1 (plus some equivalent shunt loss) in the feedback loop of an oscillator. At the point where the loop gain of the oscillator is unity, the oscillator provides exactly the required negative resistance to cancel the equivalent shunt loss of the lossy LC circuit.† In addition, most oscillators contain the necessary amplitude stabilization mechanism to ensure the self-starting of the FM generator as well as a specified level for the envelope of the output FM signal.

Although the oscillator is capable of cancelling the LC circuit loss, the Q of the LC circuit should be kept as high as possible so that, if the oscillator limits in a nonlinear fashion, the sinusoidal waveform of the output FM signal is maintained. In addition, the time constant of the envelope-limiting circuitry (cf. Section 6.8) should be fast in comparison with variations in $f(t)$ to ensure that the output envelope readjusts itself to a fixed level for each small variation in $f(t)$ and thus is independent of $f(t)$. The fast time constant also provides the continuous updating of the initial conditions of $v(t)$, thereby further ensuring close correspondence between the solutions of the exact and of the quasi-static FM differential equations.

A practical LC circuit employing a differential pair which implements the circuit model of Fig. 11.5–1(a) is shown in Fig. 11.5–2. For this circuit the capacitor current i_C is sampled by the small resistor r† (which is not loaded by the input impedance of Q_3) to develop a voltage $v_1 = i_C r$ at the base of Q_3. The amplification of the differential pair, which is directly controlled by $I_k(t)$, produces a collector current at Q_3 which simulates the time-varying controlled source $A(t)i_C$ where, on a small-signal basis,

$$A(t) = g_m(t)r = \overbrace{\frac{\alpha q r I_{k0}}{4kT}}^{A_0} + \overbrace{\frac{\alpha q r I_{k1}}{4kT}}^{A_1} f(t). \qquad (11.5\text{–}11)‡$$

Therefore, for the circuit of Fig. 11.5–2 (cf. Eq. 11.5–8),

$$\omega_0 = \frac{1}{\sqrt{L(C_0 + C)(1 + B)}} \qquad (11.5\text{–}12)$$

---

† Consider, for example, the oscillator of Fig. 6.1–6. Since $nv_t = v_i$, the controlled current source $g_m v_i$ is equal to $ng_m v(t)$. However, the voltage across the current source is v_t; thus the current source may be replaced by an equivalent negative conductance $-ng_m$. With $A_L(j\omega_0) = 1$, $-ng_m = -(G_L + n^2 G_{\text{in}})$, which is exactly the negative of the total circuit loss, $G_L + n^2 G$. Therefore, the circuit loss is canceled, leaving a pure LC network.

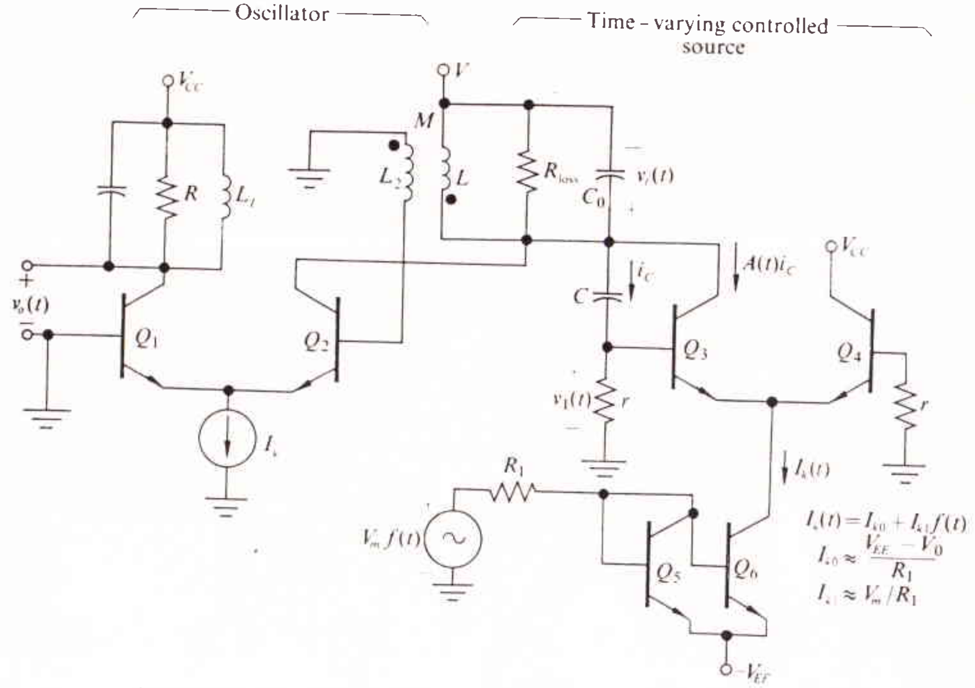


Fig. 11.5–2 Quasi-static FM generator employing a differential pair.

---

† The value of r should be sufficiently small so that, when it is transformed into an equivalent parallel resistor at ω_0, it is much greater than R_{loss} and thus is a negligible component of the total tuned-circuit loss. If this were not the case, the total tuned-circuit loss would be a function of r; however, r, like C, varies with $A(t)$. Therefore, the envelope level at which the FM wave would stabilize (which is directly related to the tuned-circuit load—see Section 6.5) would be a function of $A(t)$; this is an undesired effect.

‡ The collector current of Q_3 also contains low-frequency terms proportional to $I_k(t)$ which are shunted to ground by L and thus do not contribute to the output.

nd

$$\frac{\Delta\omega}{\omega_0} = \frac{-B(I_{k1}/I_{k0})}{2(1 + B)}, \qquad (11.5\text{–}13)$$

here $B = C\alpha qrI_{k0}/4kT(C_0 + C)$. As a result, $v_t(t)$ is given by

$$v_t(t) = V_t \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \theta_0\right] \qquad (11.5\text{–}14)$$

nd $v_o(t)$ is given by

$$v_o(t) = V_{CC} + V_0 \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \theta_0\right], \qquad (11.5\text{–}15)$$

provided that $\beta = \Delta\omega/\omega_m > 1$, $\omega_0 > 75\,\Delta\omega$, and the output-tuned circuit has sufficient bandwidth to pass the spectrum of the FM signal. Both V_t and V_0 are determined by the amplitude-limiting mechanism of the oscillator in the fashion described in Section 6.5.

To choose values for the various parameters of the time-varying controlled source, we first observe that B is a function of temperature and should be as small as possible in comparison with unity to keep ω_0 independent of T. The limit on the smallness of B is determined by the required ratio of $\Delta\omega/\omega_0$ and the ratio I_{k1}/I_{k0} (cf. Eq. 11.5–13). For a specified ratio of $\Delta\omega/\omega_0$ the minimum value of B results when I_{k1}/I_{k0} is as large as possible. The limit on the size of this ratio is 1, since $I_k(t) \geq 0$; therefore, a sensible choice would be to select $I_{k1} \approx I_{k0}$ and to select B with its minimum value of

$$B = -\frac{2\,\Delta\omega/\omega_0}{1 + 2\,\Delta\omega/\omega_0} \approx \left|\frac{2\,\Delta\omega}{\omega_0}\right|, \qquad \omega_0 \gg \Delta\omega. \qquad (11.5\text{–}16)$$

With this value of B, ω_0 simplifies to

$$\omega_0 = \frac{1}{\sqrt{L(C_0 + C)\dfrac{1}{1 + 2\,\Delta\omega/\omega_0}}} \approx \frac{1}{\sqrt{L(C_0 + C)}}, \qquad \omega_0 \gg \Delta\omega. \qquad (11.5\text{–}17)$$

We also observe that to keep Q_3 operating as a linear controlled source, Q_3 must function as a small-signal amplifier or, equivalently, $|v_1| < 10\text{ mV}$. Since $r \ll 1/\omega_0 C$ and $\omega_0 \gg \Delta\omega$, the amplitude of i_C may be approximated by $V_t\omega_0 C$; thus the amplitude of v_1 may be approximated by $V_t r\omega_0 C$. Combining the requirement of $V_t r\omega_0 C < 10\text{ mV}$ with the definition of B (and letting $\alpha \approx 1$ and $kT/q = 25\text{ mV}$), we obtain

$$\frac{\omega_0(C_0 + C)BV_t}{I_{k0}} \approx \frac{BV_tQ_T}{I_{k0}R_{\text{loss}}} < 0.1, \qquad (11.5\text{–}18)$$

where $Q_T \approx \omega_0(C_0 + C)R_{\text{loss}}$ is the Q of the circuit with $f(t) = 0$. Since B, V_t, Q_T, and R_{loss} are usually known or specified quantities, Eq. (11.5–18) determines the minimum value of I_{k0}.

As a numerical example, consider the design of an FM generator of the form shown in Fig. 11.5–2, for which $\omega_0 = 10^8$ rad/sec, $\Delta\omega = 10^6$ rad/sec, $\omega_m = 2 \times 10^5$ rad/sec, $Q_T = 20$, $V_t = 5$ V, $V_{CC} = 10$ V, and $R_{\text{loss}} = 10\text{ k}\Omega$. From Eq. (11.5–16) we obtain the optimum value of B for a temperature-insensitive center frequency, namely, $B \approx 2\,\Delta\omega/\omega_0 = 0.02$ provided that $I_{k1} \approx I_{k0}$. In addition, from Eq. (11.5–18) we find that I_{k0} must be greater than 2 mA. A choice of $I_{k1} = I_{k0}$ of 2.5 mA would be reasonable here. In addition, since $Q_T = \omega_0(C_0 + C)R_{\text{loss}}$, then $C_0 + C = 20$ pF, a value which when substituted into Eq. (11.5–15) yields $L = 5\ \mu$H.

If we choose $C_0 = C = 10$ pF (the ratio of C to C_0 may be chosen arbitrarily), then with $kT/q = 26$ mV and $\alpha \approx 1$, from the definition of B we obtain $r = 1.6\ \Omega$. Clearly r is not loaded by Q_3.

The oscillator differential pair can now be designed to yield the correct value of V_t. Usually the oscillator is designed to stabilize with a large value of x such that, even if small amplitude variations exist at $v(t)$, $v_0(t)$ has a constant amplitude. This effect follows from the fact that with large values of x the output of Q_1 is a limited version of the input to Q_2.

To cause the oscillator to stabilize at $x = 10$ we choose M such that

$$\frac{M}{L} = \frac{kTx}{qV_t} = 0.05 \qquad \text{or} \qquad M = 0.25\ \mu\text{H}$$

and I_k such that

$$1 = A_L(j\omega_0) \approx \frac{MG_m(10)}{L}R_{\text{loss}}.$$

Thus $G_m(10)$ must equal $1/500\ \Omega$ and from Fig. 4.6–5 g_m must equal $1/130\ \Omega$. However, since $g_m = qI_k/4kT$, $I_k \approx 800\ \mu$A.

Now if $R_L = 6\text{ k}\Omega$ and the output-tuned circuit bandwidth is sufficient for the FM signal ($\text{BW}/2\,\Delta\omega > 1.6$ with $\beta = 5$, or $\text{BW} > 3.2 \times 10^6$ rad/sec), then

$$v_o = (10\text{ V}) - (6\text{ V})\cos\left[10^8 t + 10^6 \int^t f(\theta)\, d\theta + \theta_0\right].$$

If we choose $\text{BW} = 5 \times 10^6$, with the result that Q_T for the output-tuned circuit is 20, then $L_L = 3\ \mu$H and $C_L = 33.3$ pF.

An alternative circuit for obtaining the time-varying controlled current source required to implement the circuit model of Fig. 11.5–1(a) is shown in Fig. 11.5–3. For this circuit, which employs a single transistor, $A(t)$ is given by

$$A(t) = g_m(t)r = \overbrace{\frac{\alpha qrI_{E0}}{kT}}^{A_0} + \overbrace{\frac{\alpha qrI_{E1}}{kT}}^{A_1} f(t),$$

provided that r is sufficiently small so that the transistor is operated in a small-signal fashion and that the impedance of C_E at ω_0 is small in comparison with the minimum value of $g_{\text{in}}(t)$ so that v_1 appears across the transistor junction. The impedance of C_E at ω_m should also be large in comparison with the maximum value of $g_m(t)$ so

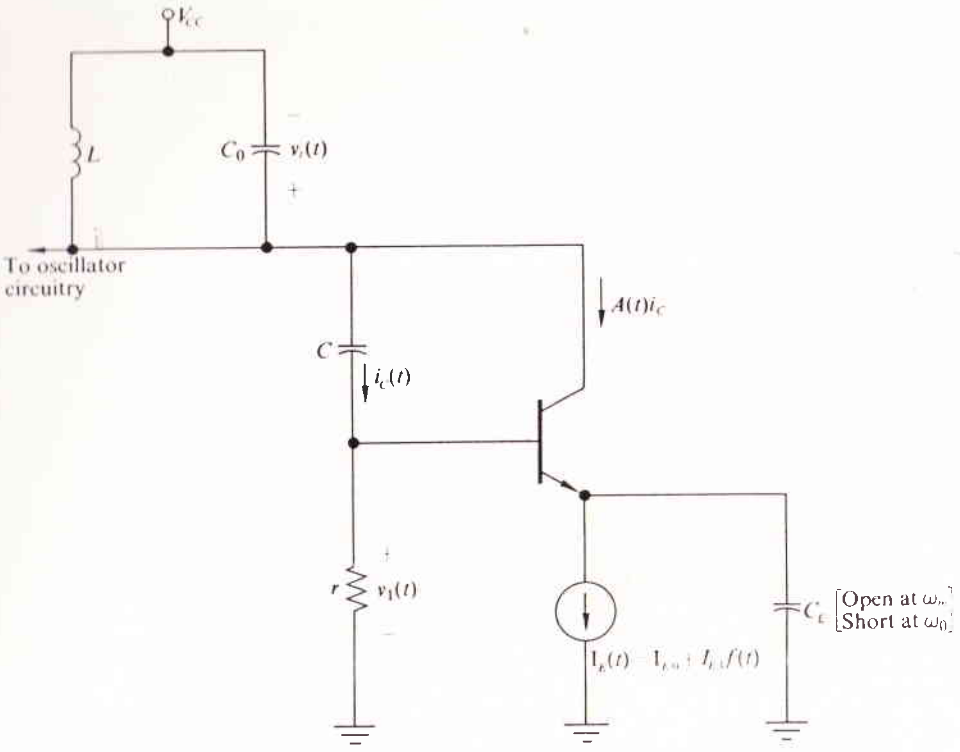


Fig. 11.5–3 Implementation of time-varying controlled source with a single transistor.

that $I_E(t)$ flows entirely through the transistor. This latter requirement can be implemented only if ω_0 exceeds ω_m by a factor of 1000 or more.

For example, if $I_{E1} = 0.9I_{E0}$ (which is desirable, as we have seen, to keep ω_0 independent of T), then

$$\frac{g_{\text{in}_{\max}}}{g_{\text{in}_{\min}}} = 19;$$

therefore, if

$$\omega_0 C_E > 10 g_{\text{in}_{\max}} \qquad \text{and} \qquad \omega_m C_E < \frac{g_{\text{in}_{\min}}}{10},$$

then ω_0/ω_m must be greater than 1900 to perform comparably with the differential pair implementation of the time-varying controlled source. Consequently, only when extremely wide separations exist between ω_0 and ω_m can the single transistor be used in place of the differential pair as the time-varying controlled source. However, when the circuit of Fig. 11.5–3 is employed, the same design criteria are used as were used for the circuit of Fig. 11.5–2. Still other variations on the implementation of the time-varying controlled current source required for the circuit of Fig. 11.5–1(a) will be explored in the problems at the end of the chapter.

The circuit model of Fig. 11.5–1(b) can be implemented in exactly the same fashion as the model of Fig. 11.5–1(a) except that an inductor L replaces the capacitor C. With the inductor, however, an additional coupling capacitor must be inserted in series with L to preserve the bias conditions of the circuits of Figs. 11.5–2 and 11.5–3. This additional large capacitor makes the inductive case less attractive than the capacitive case. A similar argument applies to the circuit model of Fig. 11.5–1(c).

A time-varying capacitance can also be obtained from any device whose stored charge q is a nonlinear function of the applied voltage v_D. Any reverse-biased PN junction has this property. For example, if a device has the relationship $q = g(v_D)$, where $g(v_D)$ represents a nonlinear function of v_D, then the device current is given by

$$i_D = \frac{dq}{dt} = \frac{dg(v_D)}{dv_D}\dot{v}_D = C(v_D)\dot{v}_D, \tag{11.5–19}$$

where $C(v_D) = dg(v_D)/dv_D$. If a high-frequency developed voltage $v(t)$ and a low-frequency bias and controlling voltage $V_{CC} + V_C(t)$ are superimposed across the device as shown in Fig. 11.5–4, then $i_D(t)$ can be written as

$$\begin{aligned} i_D(t) &= C[-V_{CC} - V_C(t) + v(t)][\dot{v}(t) - \dot{V}_C(t)] \\ &\approx C[-V_{CC} - V_C(t) + v(t)]\dot{v}(t). \end{aligned} \tag{11.5–20}$$

[It is apparent that $\dot{V}_C(t)$ is negligible in comparison with $\dot{v}(t)$ because of the large frequency separation between these two voltages.]

In addition, if it is assumed that $V_C(t) \ll V_{CC}$ and $v(t) \ll V_{CC}$, then Eq. (11.5–18) can be approximated by its first two power-series terms in the form

$$i_D = \overbrace{C(-V_{CC})}^{C_0}\dot{v}(t) + \frac{\partial C(-V_{CC})}{\partial V_{CC}}v(t)\dot{v}(t) \overbrace{- \frac{\partial C(-V_{CC})}{\partial V_{CC}}V_C(t)}^{C(t)}\dot{v}(t). \tag{11.5–21}$$

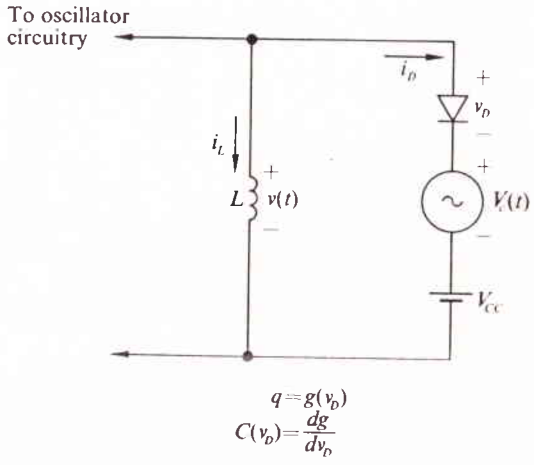


Fig. 11.5–4 Nonlinear capacitor employed in quasi-static FM generator.

nce the product $v(t)\ddot{v}(t)$ produces terms in the vicinity of dc and $2\omega_0$ which are not assed by the LC circuit of Fig. 11.5–4, the second component of i_D may be neglected the complete differential equation for $v(t)$:

$$0 = v(t) - \frac{2\dot{\omega}_i(t)}{\omega_i(t)^3}\dot{v}(t) + \frac{\ddot{v}(t)}{\omega_i(t)^2}, \tag{11.5–22}$$

here $\omega_i(t) = 1/\sqrt{L[C_0 + C(t)]}$. Clearly Eq. (11.5–20) is identical with Eq. (11.5–1b) nd thus the circuit of Fig. 11.5–4 may be employed as a quasi-static FM generator. A pecific application of this circuit is explored in the problems at the end of the chapter. n practical situations one increases both the linearity and the deviation sensitivity y using two diodes in a push-pull arrangement.

Special hyper-abrupt junctions are available in which $C(t)$ varies as the square of ne signal voltage. With such units the static restrictions on maximum deviation be- ore excessive distortion may be somewhat relaxed.

In the older literature the circuit ideas described in this section were known as reactance tube" and "varactor tuned" FM generators. Both types of circuits have een in use for over 40 years. The static restrictions on the vacuum tube versions ere well described more than 25 years ago. The dynamic restrictions and the elationship of these generators to the exact FM equation are not known to have been resented previously.

1.6 TRIANGULAR-WAVE FREQUENCY MODULATION

he generation and nonmemory waveshaping of an FM triangular wave is the most ractical method of generating a sinusoidal FM signal which has no theoretical estrictions on either its frequency deviation $\Delta\omega$ or its maximum modulation fre- uency ω_m. (The restriction $\Delta\omega < \omega_0$, to avoid full-wave rectification of the in- stantaneous frequency, does, of course, apply.) In this section we shall study the roblem of FM triangular wave generation by first presenting a general block diagram f a triangular wave generator and then implementing the block diagram with physical ircuitry. We shall also consider the implementation of a waveshaping circuit which onverts the FM triangular wave into a sinusoidal FM signal.

As was pointed out in Section 11.2, an FM triangular wave $v_T(t)$ is a symmetric iangular voltage which is periodic (of period $T = 2\pi/\omega_0$) when plotted vs.

$$\tau(t) = t + \frac{\Delta\omega}{\omega_0}\int^t f(\theta)\, d\theta$$

f. Fig. 11.2–6). We also recall from Section 11.2 that such a waveform has the nstantaneous frequency

$$\omega_i(t) = \omega_0 + \Delta\omega f(t). \tag{11.6–1}$$

Figure 11.6–1 illustrates the block diagram of a system which produces a periodic iangular wave in $\tau(t)$ of peak-to-peak amplitude $2V_2$ and thus an FM triangular wave

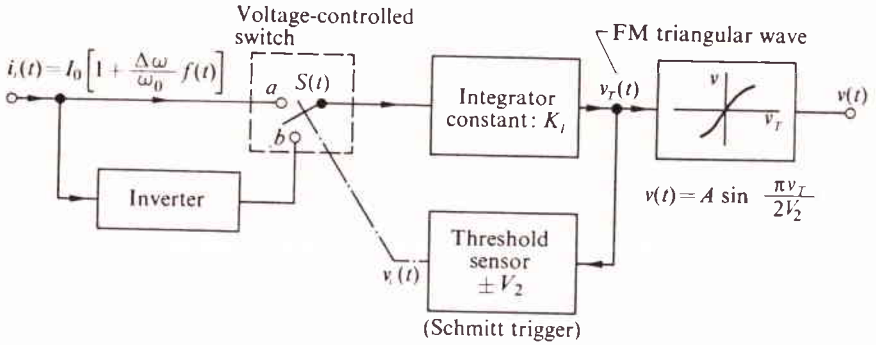

Voltage-controlled switch
FM triangular wave
$i_i(t) = I_0\left[1 + \frac{\Delta\omega}{\omega_0} f(t)\right]$
a
S(t)
b
Integrator constant: K_I
$v_T(t)$
v
v_T
v(t)
Inverter
$v(t) = A \sin \frac{\pi v_T}{2V_2}$
Threshold sensor $\pm V_2$
$v_s(t)$
(Schmitt trigger)

Fig. 11.6–1 Block diagram of an FM triangular wave generator.

of the same peak-to-peak amplitude in t. The system generates an alternating sequence of positively and negatively sloping ramps in $\tau(t)$ by integrating $+i_i(t)$ and $-i_i(t)$ respectively, where

$$i_i(t) = I_0\left[1 + \frac{\Delta\omega}{\omega_0} f(t)\right].$$

To obtain the correct peak-to-peak amplitude for the alternating sequence of ramps at the integrator output, a threshold sensor switches the integrator input from $i_i(t)$ to $-i_i(t)$ when the integrator output reaches V_2 and switches the integrator input back to $i_i(t)$ when the integrator output reaches $-V_2$.

More specifically, when the input voltage to the threshold sensor reaches a value of V_2, a step of voltage appears at the output, causing the voltage-controlled $S(t)$ to switch from contact a to contact b. Similarly, when the input voltage to the sensor reaches $-V_2$, a step of voltage opposite to the original step causes $S(t)$ to return from contact b to contact a. Figure 11.6–2 illustrates the input-output characteristic of the threshold sensor required to operate the voltage-controlled switch, as well as the corresponding plot of $S(t)$ vs. $v_T(t)$.

To obtain a more quantitative expression for $v_T(t)$ assume, at $t = t_i$, $v_T(t_i) = V_2$ such that $S(t)$ switches from a to b. Hence, for $t > t_i$,

$$-i_i(t) = -I_0\left[1 + \frac{\Delta\omega}{\omega_0} f(t)\right]$$

is applied to the integrator of Fig. 11.6–1 and $v_T(t)$ is given by

$$\begin{aligned} v_T(t) &= V_2 - K_I \int_{t_i}^{t} i_i(\theta)\, d\theta \\ &= V_2 - K_I I_0\left[(t - t_i) + \frac{\Delta\omega}{\omega_0}\int_{t_i}^{t} f(\theta)\, d\theta\right] \qquad (11.6–2) \\ &= V_2 - K_I I_0[\tau(t) - \tau(t_i)], \qquad t_i < t \le t_{i+1}, \end{aligned}$$

where K_I is the integrator constant. Clearly, for $t > t_i$, $v_T(t)$ is a negative ramp with

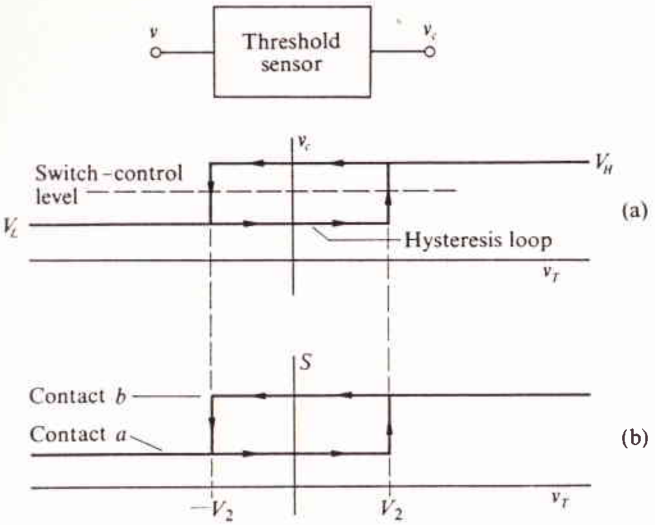


Fig. 11.6–2 (a) Plot of v_c vs. v_T for the threshold sensor. (b) Plot of switch position vs. v_T for the threshold sensor–controlled switch combination.

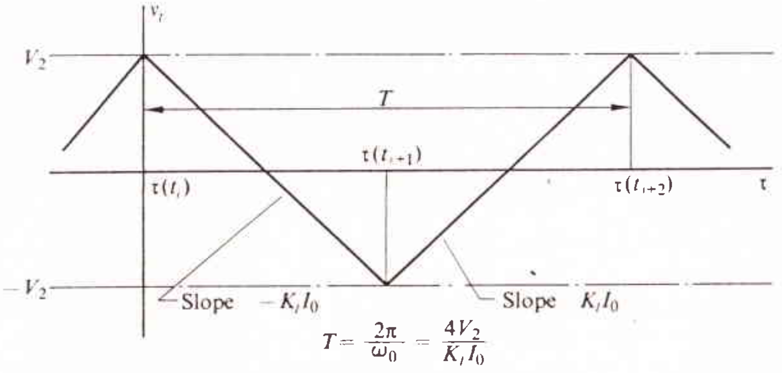


Fig. 11.6–3 Plot of $v_T(t)$ vs. $\tau(t)$.

slope $-K_I I_0$ in τ, as shown in Fig. 11.6–3. The ramp reaches $-V_2$ at $t = t_{i+1}$, at which time $S(t)$ switches to a. For $t > t_{i+1}$, $i_i(t)$ is applied to the integrator and $v_T(t)$ takes the form

$$v_T(t) = -V_2 + K_I I_0[\tau(t) - \tau(t_{i+1})], \tag{11.6–3}$$

which is a positive ramp of slope $K_I I_0$ in τ. When $v_T(t)$ again reaches V_2 at $t = t_{i+2}$ the cycle repeats.

It is now quite apparent from Fig. 11.6–3 that $v_T(t)$ is indeed a periodic triangular wave in τ with a period

$$T = \frac{4V_2}{K_I I_0} \tag{11.6–4}$$

and a peak-to-peak amplitude $2V_2$. Thus if $v_T(t)$ is placed through a nonlinear device with a transfer function given by

$$v = A \sin \frac{\pi v_T}{2V_2}, \qquad -V_2 \le v_T \le V_2, \tag{11.6–5}$$

the nonlinear device output $v(t)$ is the desired FM signal

$$\begin{aligned} v(t) &= A \cos \omega_0 \tau(t) \\ &= A \cos \left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta + \theta_0 \right], \end{aligned} \tag{11.6–6}$$

where

$$\omega_0 = \frac{2\pi}{T} = \frac{\pi K_I I_0}{2V_2}$$

and θ_0 is an arbitrary phase angle which may be taken as zero with the correct choice of time origin.

It is of interest to note that the voltage at the output of the threshold sensor sits at one voltage level during the positively sloping portions of $v_T(t)$ and sits at another voltage level during the negatively sloping portions of $v_T(t)$. Hence, when plotted versus $\tau(t)$, the threshold sensor output is a periodic square wave with the same period as $v_T(t)$. Therefore, the threshold sensor output vs. time is an FM square wave. Consequently it is possible to obtain from a single network an FM triangular wave, an FM square wave, and an FM sinusoidal signal. (Note the 90° phase shift between the square wave and the two other waveforms.) However, if only the FM square wave is desired, a much simpler technique for generating it is available. Such a technique is explored in Section 11.7. The ability to generate several different types of FM signals simultaneously makes the triangular FM generator quite valuable as a piece of laboratory test equipment.

The block diagram of Fig. 11.6–1 can be physically implemented by selecting circuits which perform the function of each block and interconnecting them. Although there are many possible circuits for each of the blocks, we shall select only one representative circuit in each case.

Threshold Sensor (Large-Hysteresis Schmitt Trigger Circuit)

The desired threshold sensor characteristic is given in Fig. 11.6–2. For high-frequency operation the distance between the two switching points should be 4 V or more. Such an amplitude will allow the triangular wave to drive the diode waveshaping network without further amplification. This direct drive is highly desirable, since to linearly amplify a triangular wave requires an extremely wideband amplifier.

Thus the basic problem in the Schmitt trigger design in this case is to maximize the hysteresis of the circuit. (In the more normal use of the circuit, one minimizes the hysteresis.) A possible Schmitt trigger circuit is illustrated in Fig. 11.6–4.

If Q_2 is allowed to move only between the active region and cutoff, then the output is always isolated from the rest of the circuit. Also, keeping Q_2 active will avoid the

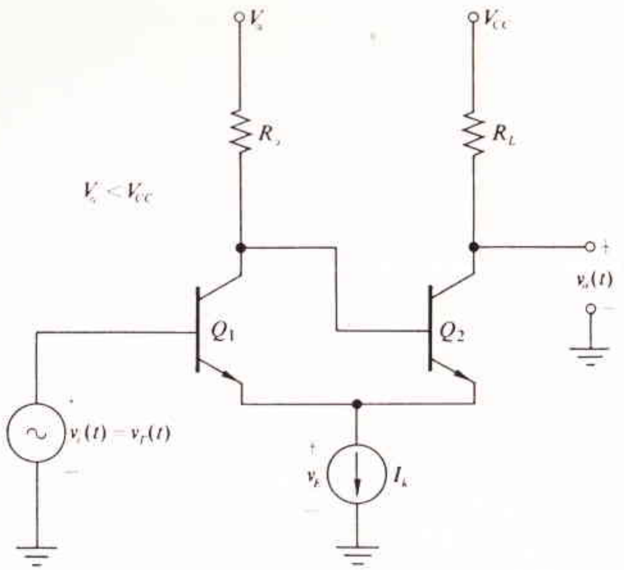


Fig. 11.6–4 Schmitt trigger circuit.

undesirable circuit "delay" that would be caused by the time necessary for Q_2 to come out of saturation.

The regeneration feedback path is supplied by the two transistors when both are active. If $v_i(t)$ is initially large and negative, then Q_1 will be off and Q_2 on:

$$v_E = [V_x - (1 - \alpha_2)R_x I_k] - V_0.$$

When $v_i(t)$ increases to the neighborhood of the bracketed term in v_E, then both transistors become active; and if the loop gain exceeds unity, regenerative switching occurs which leaves Q_1 on and Q_2 off. Increasing $v_i(t)$ further has no effect on the output voltage. If one starts to decrease $v_i(t)$, then reverse switching will not occur until Q_1 comes out of saturation far enough so that a positive base-emitter voltage may be applied to Q_2. This requires $v_i(t) \approx V_x - \alpha_1 R_x I_k$. Again, regenerative action and sudden switching to the initial state will occur.†

If the two input switching levels are denoted by v_{il} and v_{iu} for lower and upper values respectively, then

$$v_{il} \approx V_x - \alpha_1 R_x I_k \qquad \text{and} \qquad v_{iu} \approx V_x - (1 - \alpha_2)R_x I_k.$$

Hence the peak-to-peak value of the triangular wave is

$$V_{pp} = -I_k R_x[\alpha_1 - (1 - \alpha_2)] \approx 0.95 I_k R_x.$$

---

† For a more detailed description of the Schmitt trigger circuit, see J. Millman and H. Taub *Pulse, Digital, and Switching Waveforms*, McGraw-Hill, New York (1965), pp. 389–402.

The output voltage of the circuit switches between V_{CC} and $V_{CC} - \alpha_2 R_L I_k$; hence the output switching amplitude is $\alpha_2 R_L I_k$. In addition, to keep Q_2 in the active region,

$$V_{CC} - \alpha_2 R_L I_k > V_x - (1 - \alpha_2) I_k R_x$$

or, as a conservative approximation,

$$V_{CC} > V_x + I_k R_L.$$

Figure 11.6–5 illustrates these characteristics.

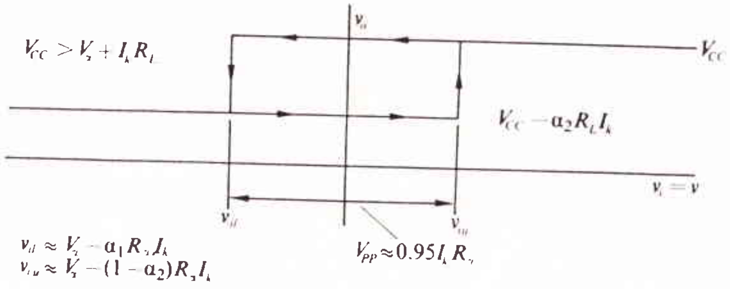

v_o
$V_{CC} > V_x + I_k R_L$
V_{CC}
$V_{CC} - \alpha_2 R_L I_k$
$v_i = v$
v_{il}
v_{iu}
$v_{il} \approx V_x - \alpha_1 R_x I_k$
$v_{iu} \approx V_x - (1 - \alpha_2) R_x I_k$
$V_{PP} \approx 0.95 I_k R_x$

Fig. 11.6–5 Schmitt trigger transfer characteristic.

For high-frequency operation one tries to keep the resistance values low (below 1 kΩ); hence for large output and input voltages the current levels are appreciable—perhaps 10 mA or more. Thus Q_2 must be checked for excessive dissipation.

Voltage-Controlled Switch, Inverter, and Integrator

The remaining portion of the circuit may be implemented by the circuit of Fig. 11.6–6. The voltage-controlled switching is accomplished by the differential pair (Q_1 and Q_2), and the capacitor C is employed as the integrator. Transistor Q_3 functions as the inverter.

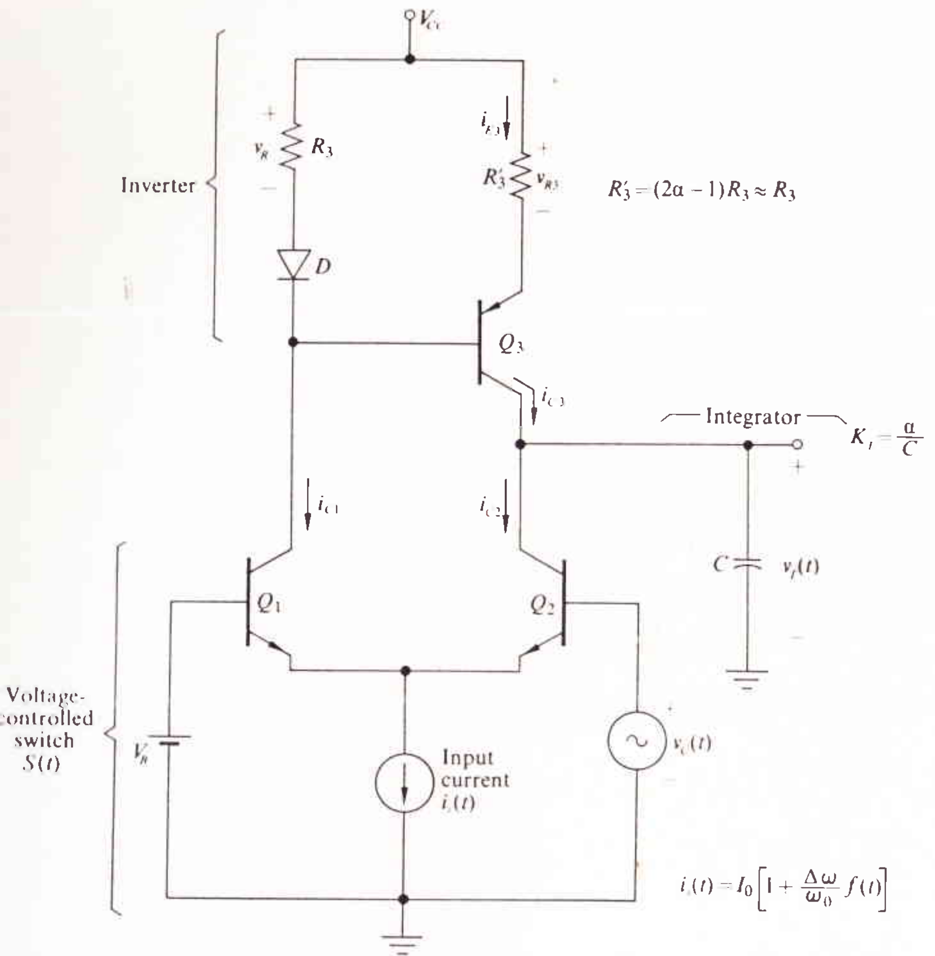
It is apparent that with $v_C(t)$ less than V_B by several times kT/q (cf. Fig. 4.6–2), $i_i(t)$ flows entirely out of Q_1, with the result that the collector current of Q_1 has the value $i_{C1} = \alpha i_i(t)$ while $i_{C2} = 0$. The collector current of Q_1, less $(1 - \alpha)i_{E3}$, flows through R_3 to develop a voltage

$$v_{R3} = [\alpha_1 i_i - (1 - \alpha_3) i_{E3}] R_3. \tag{11.6–7}$$

With the assumption that the junction voltage across D equals the emitter-base voltage of Q_3 (which is true if D and Q_3 are integrated on the same chip and have the same geometry and if $i_D \approx i_{E3}$), $v_{R3'} = v_{R3} = R'_3 i_{E3}$; thus

$$i_{E3} = \frac{\alpha_1 i_i}{1 - \alpha_3 + \dfrac{R'_3}{R_3}}. \tag{11.6–8}$$

Now $i_{C3} = \alpha_3 i_{E3}$ flows into the integrating capacitor.



ig. 11.6-6 Practical implementation of the voltage-controlled switch, the inverter, and the itegrator of Fig. 11.6–1.

On the other hand, with $v_c(t)$ greater than V_B by several times kT/q, $i_i(t)$ flows en-rely out of Q_2, $i_{E3} = i_{C3} = i_{C1} = 0$, and $i_{C2} = \alpha_2 i_i$ flows out of the integrating ipacitor. Consequently, if the alphas are equal and the ratio R_3'/R_3 is adjusted to qual $2\alpha - 1$ so that $i_{E3} = i_i$ in Eq. (11.6–8), then as v_c experiences a step increase from veral times kT/q below V_B to several times kT/q above V_B, the integrator current vitches from $\alpha i_i(t)$ to $-\alpha i_i(t)$; this is the desired effect for the voltage-controlled vitch-inverter combination.

We note that, when $\alpha i_i(t)$ flows into the integrator, the output voltage $v_T(t)$ is ven by

$$v_T(t) = \frac{\alpha}{C} \int^t i_i(\tau)\, d\tau; \tag{11.6–9}$$

therefore, the integrator constant K_I for the circuit of Fig. 11.6–6 has the value

$$K_I = \frac{\alpha}{C}. \tag{11.6–10}$$

We also note that, to keep Q_2 from saturating $v_c(t)_{\max}$ must be less than $v_T(t)_{\min} + V_0$; to keep Q_3 from saturating, $V_{CC} - i_i(t)_{\max} R_3$ must be greater than $v_T(t)_{\max}$.

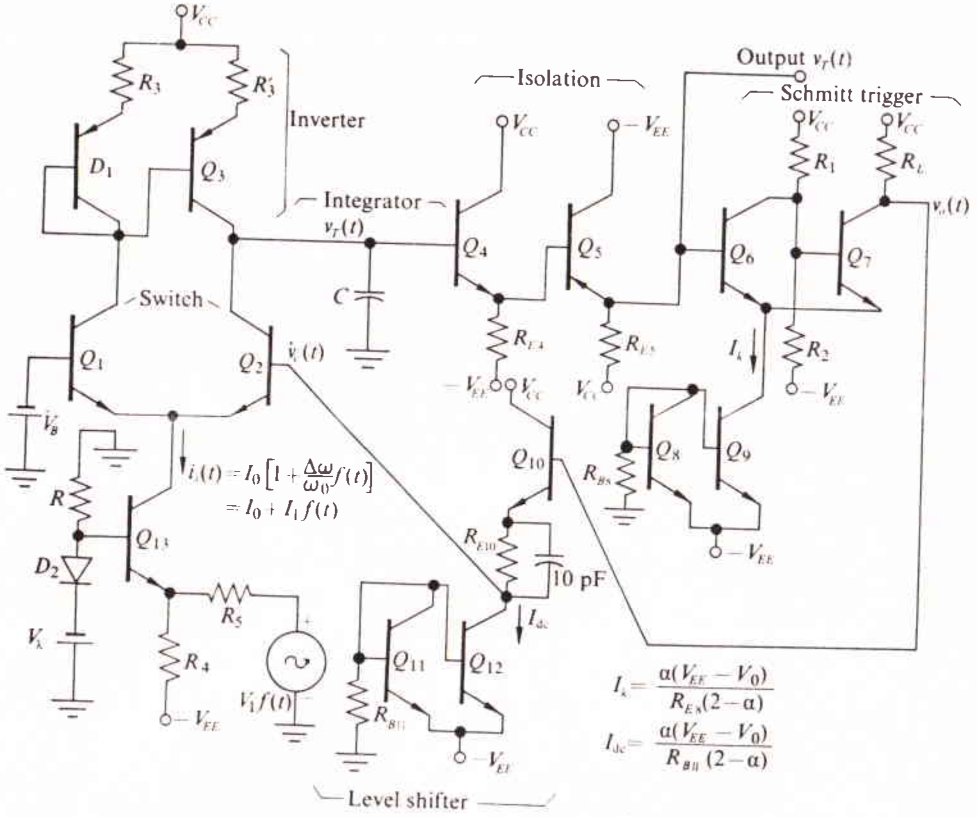


Fig. 11.6-7 Complete FM triangular wave generator.

Figure 11.6–7 illustrates the complete circuit for the FM triangular wave generator which combines the threshold sensor of Fig. 11.6–4 with the circuit of Fig. 11.6–6. Two emitter followers with compensating base-emitter junction biases are employed to isolate the integrating capacitor from the input to the Schmitt trigger. In addition, a dc-level shifter is incorporated at the Schmitt trigger output, with the result that

$$v_c(t) = v_o(t) - V_0 - I_{dc} R_{E10}. \tag{11.6–11}$$

This shift permits $v_c(t)_{\max}$ to be kept less than $v_T(t)_{\min} + V_0$, as is required to keep Q_2 from saturating. For example, if $v_T(t)_{\min} = -V_2 = -3$ V and $V_{CC} = 10$ V, then a choice of $I_{dc}R_{E10}$ of 13 V would prevent Q_2 from saturating. In addition, if $V_{CC} - \alpha R_L I_k$ for the Schmitt trigger is 5 V, then $v_c(t)$ switches between $-(3\text{ V}) - V_0$ and $-(8\text{ V}) - V_0$. For this case V_B should be chosen between the two limits of $v_c(t)$; $V_B = -5$ V would be reasonable. The 10 pF capacitor across R_{E10} couples the fast rise and fall times of $v_o(t)$ to $v_c(t)$.

If for the FM triangular wave generator of Fig. 11.6–7 the Schmitt trigger switches states when its input reaches $\pm V_2$ and if $R'_3 = (2\alpha - 1)R_3$ so that $K_I = \alpha/C$, then the carrier frequency is given by (cf. Eq. 11.6–6)

$$\omega_0 = \frac{\pi K_I I_0}{2V_2} = \frac{\pi\alpha I_0}{2V_2 C}, \tag{11.6–12}$$

which is inversely proportional to C. Hence, for any given values of I_0 and V_2, C may be chosen to yield the desired center frequency for the FM triangular wave. For example, if $I_0 = 3$ mA and $V_2 = 3$ V, $C = 157$ pF yields a carrier frequency of 10^7 rad/sec.

To obtain a specified frequency deviation, we first note that $i_i(t)$ may be written in the form

$$i_i(t) = I_0 + I_1 f(t), \tag{11.6–13}$$

where $I_1 = I_0\,\Delta\omega/\omega_0$. We then choose I_1 to obtain the desired $\Delta\omega$. For example, if $I_0 = 3$ mA and $\omega_0 = 10^7$ rad/sec, a frequency deviation equal to one-half the carrier frequency may be obtained by choosing $I_1 = I_0/2 = 1.5$ mA.

The currents I_0 and I_1 are developed by Q_{13}. The diode D_2 in the base circuit of Q_{13} compensates for the base-emitter turn-on bias so that

$$I_0 = \alpha\left(\frac{V_{EE} - V_\lambda}{R_4} - \frac{V_\lambda}{R_5}\right) \tag{11.6–14}$$

and

$$I_1 = \frac{\alpha V_1}{R_5}. \tag{11.6–15}$$

It is apparent that V_λ must be selected below $-V_B - V_0$ to keep Q_{13} from saturating. For $V_B = -5$ V, a reasonable value for V_λ would be 6 V.

Now if $V_1 = 6$ V, $V_\lambda = 6$ V, $V_{EE} = 10$ V, and $\alpha \approx 1$, and we desire $I_0 = 3$ mA and $I_1 = 1.5$ mA to achieve $\omega_0 = 10^7$ rad/sec and $\Delta\omega = 10^7/2$ rad/sec, then from Eq. (11.6–15)

$$R_5 \approx \frac{V_1}{I_1} = 4\text{ k}\Omega$$

and from Eq. (11.6–14)

$$R_4 \approx \frac{V_{EE} - V_\lambda}{V_\lambda/R_5 + I_0} = 890\ \Omega.$$

To convert an FM triangular wave to a sinusoidal FM signal, a nonlinear network with the transfer characteristic

$$v = A \sin \frac{\pi v_T}{2V_2} \tag{11.6–16}$$

is required.† Figure 11.6–8 illustrates this characteristic for the case where $A = 1$ V and $V_2 = 2$ V. In addition, Fig. 11.6–8 includes the break points and the slope of a possible seven-segment piecewise-linear approximation to the nonlinear characteristic. More segments can be added if a closer approximation is desired.

Figure 11.6–9 illustrates two possible diode waveshaping networks which physically realize the piecewise-linear characteristic shown in Fig. 11.6–8. The voltage sources in the network are chosen to correspond to the breakpoints of Fig. 11.6–8,

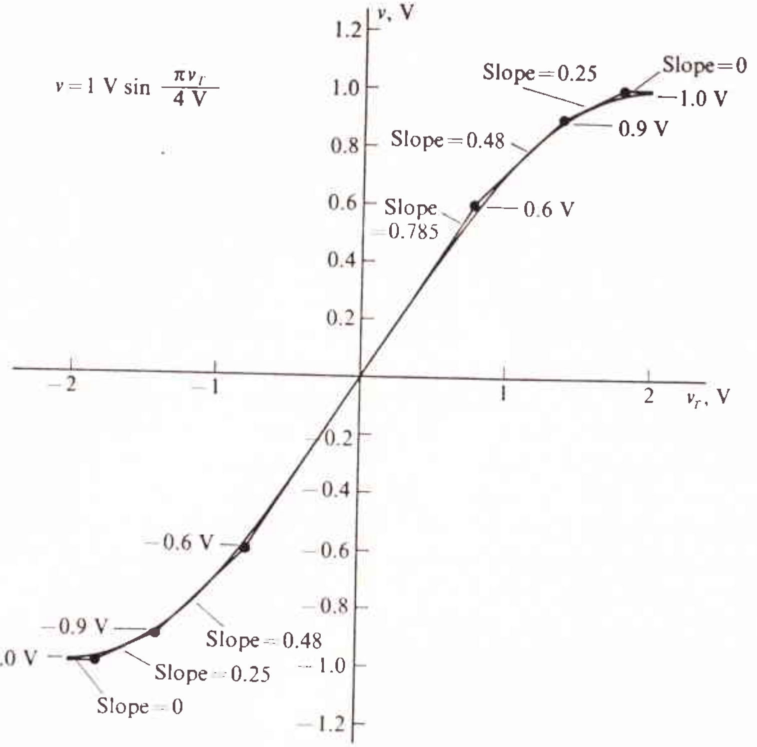
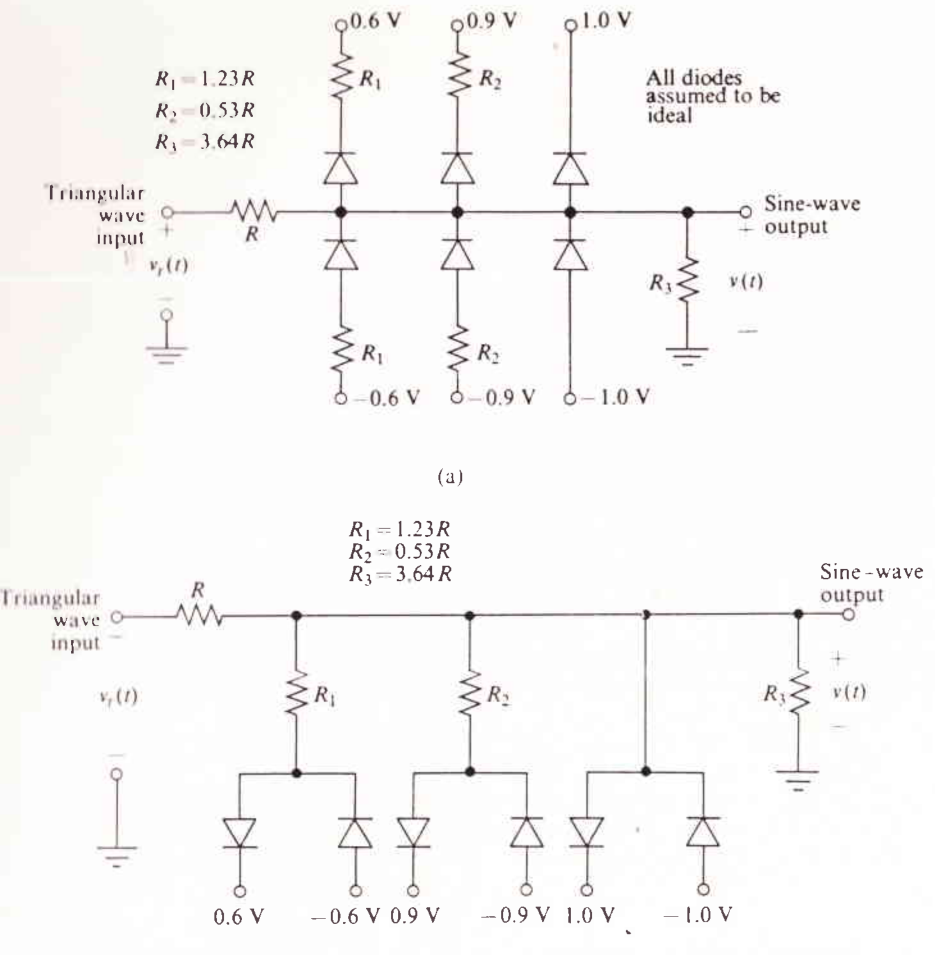


Fig. 11.6–8 Plot of $v = (1\text{ V}) \sin [\pi v_T/(4\text{ V})]$ vs. v_T plus a seven-segment piecewise-linear approximation of the characteristic.

---

† See L. Strauss, *Wave Generation and Shaping*, McGraw-Hill, New York (1960). Section 2–8 and Problem 2–22 discuss and illustrate a diode waveshaping network. See also G. Klein, "Accurate Triangle–Sine Converter," *ISSCC Digest of Technical Papers (Philadelphia)*, pp. 120–121 (Feb. 1967).



(a)

(b)

g. **11.6–9** (a) Diode waveshaping network to realize the piecewise-linear approximation of
g. 11.6–8. (b) More practical version of a diode waveshaping network to realize the approxi-
ation of Fig. 11.6–8.

hile the resistors R_1, R_2, and R_3 are selected to give the correct slope in each region.
or example, $R_3/(R_3 + R)$ must equal 0.785 to produce the initial slope of the piece-
se-linear characteristic; therefore, $R_3 = 3.64$ R. Similarly,

$$\frac{R_1 \| R_3}{(R_1 \| R_3) + R} = 0.48 \quad \text{and} \quad \frac{R_1 \| R_2 \| R_3}{(R_1 \| R_2 \| R_3) + R} = 0.25,$$

ith the result that $R_1 = 1.23R$ and $R_2 = 0.53R$. If more segments are employed in
e piecewise-linear approximation of the nonlinear characteristic, several more
ode–resistor–voltage-source branches would be required in the circuit of Fig.

11.6–9. When physical diodes are used in place of the ideal diodes of Fig. 11.6–9, each dc source should be reduced in magnitude by V_0 to account for the turn-on bias of the diodes.

Since the key to low distortion in such a network is symmetry it is apparent that the circuit of Fig. 11.6–9(b) is much to be preferred. This is true because it avoids the necessity of matching pairs of resistors for R_1 and R_2, etc., as is the case for circuits of the type represented by Fig. 11.6–9(a)

With well-matched commercial silicon diodes (1N914B's, for example) and commercial 1% resistors, a five-breakpoint (per half-cycle) network can yield total harmonic distortions of 56 dB below the fundamental. When "hot carrier" diodes ($V_0 \approx 0.4$ V) are used to minimize charge storage problems, then carrier frequencies of up to 30 MHz or more can be handled.

11.7 PRACTICAL SQUARE-WAVE FREQUENCY MODULATION

Square-wave frequency modulation followed by narrowband filtering is employed in many practical situations. In particular, it is applicable in those situations where the deviation ratio $D = \Delta\omega/\omega_0$ must be greater than the few percent possible with the quasi-static FM modulators and yet not so great that D determines a point above the top curve of Fig. 11.2–3 and thus prevents the extraction of the fundamental FM wave by filtering.† In general, this range of deviation ratios usually arises when a large number of reasonably low-frequency carriers are to be modulated in frequency by independent information signals, so that they can be multiplexed in frequency, and then employed collectively as the modulation for a high-frequency FM or AM signal. Many systems which must transmit a large quantity of independent data or a large number of voice channels do so in this fashion.

Although an FM square wave can be obtained as a by-product of generating an FM triangular wave (as we saw in Section 11.6), an FM square wave can be generated directly and more simply by controlling the frequency of a symmetric astable multivibrator having the basic form shown in Fig. 11.7–1. Each amplifier in the multivibrator of Fig. 11.7–1 is assumed to have the input, output, and transfer characteristics shown in Fig. 11.7–2. These characteristics closely approximate those of integrated inverters which incorporate an output buffer amplifier.

Before determining how the astable multivibrator performs as an FM generator, let us review its basic operation for the case where $I(t) = I_0$ (a constant).‡ For this case, the waveforms of v_{i1}, v_{i2}, v_{o1}, and v_{o2} are shown in Fig. 11.7–3. Referring to the waveforms, we observe that, for $t < 0$, $v_{o1} = 0$ and $v_{i2} < 0$; thus the input to amplifier 2 appears as an open circuit (amplifier 2 is cut off), while the output of amplifier 1

---

† Since a square wave contains only odd harmonics of its fundamental frequency, the filter following the FM square wave generator must only be capable of separating the FM wave centered at ω_0 from the FM wave centered at $3\omega_0$.

‡ In this treatment we are assuming that the reader has some familiarity with the basic operation of multivibrators. If this is not the case, he should consult either J. Millman and H. Taub, *Pulse, Digital, and Switching Waveforms*, McGraw-Hill, New York (1965), Chapter 11, or L. Strauss, *Wave Generation and Shaping*, 2nd edition, McGraw-Hill, New York (1970), pp. 385–389.

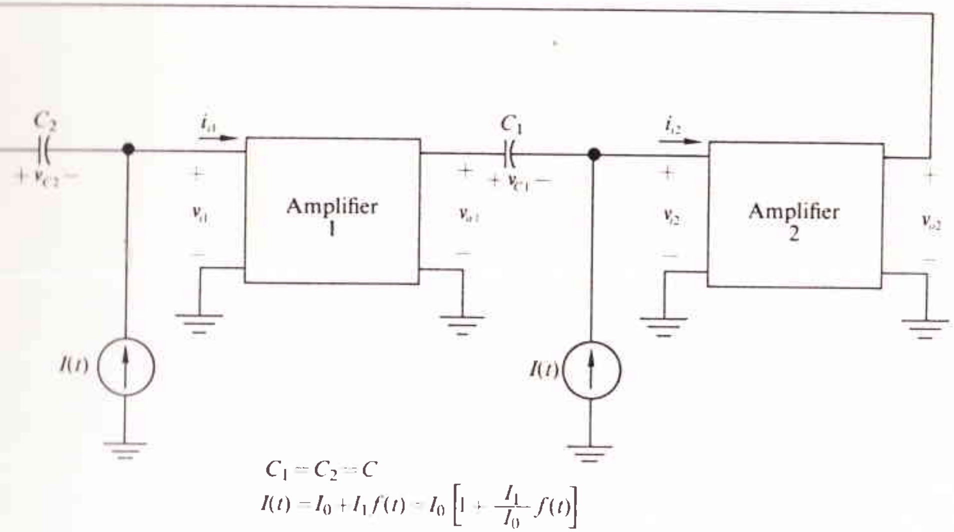


Fig. 11.7-1 FM square-wave generator employing a stable multivibrator.

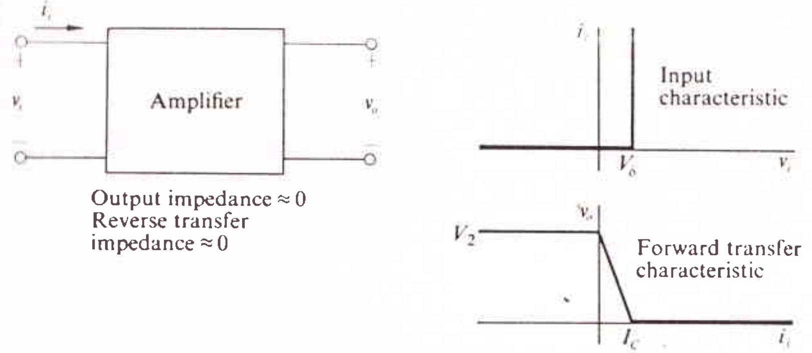


Fig. 11.7-2 Input and transfer characteristics for the amplifiers in Fig. 11.7-1.

ppears as a short circuit (amplifier 1 is saturated). Therefore, I_0 flows entirely into $_1$, causing $v_{i2}(t)$ to increase as a linear ramp with slope I_0/C. However, at $t = 0$, hen $v_{i2}(t) = V_\delta$, the input to amplifier 2 appears as an incremental short circuit f. Fig. 11.7-2) causing I_0 to flow entirely into amplifier 2 instead of C_1. If $I_0 > I$ which we assume it is), then $v_{o2}(t)$ drops instantly to zero, with the result that $v_{i1}(t)$ ops an equal amount V_2, since $v_{C2}(t)$ cannot change instantaneously. The decrease v_{i1} below V_δ causes i_{i1} to equal 0 and v_{o1} to rise rapidly to V_2. (Even though connected directly across v_{o1}, the small assumed output impedance permits be charged almost instantaneously.) At this point I_0 flows into C_2, causing increase as a ramp with slope I_0/C while $v_{o1} = V_2$, $v_{i2} = V_\delta$, and $v_{o2} = 0$. When, at $T/2$, $v_{i1} = V_\delta$, the circuit switches to its original state and the process continues

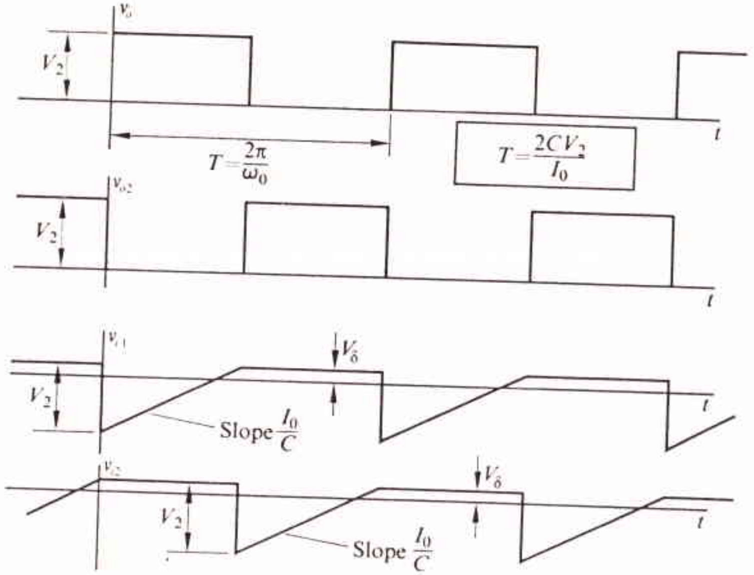


Fig. 11.7-3 Waveforms for the circuit of Fig. 11.7-1 with $I(t) = I_0$.

From the waveform of v_{i1} or v_{i2} it is apparent that the slope I_0/C times $T/2$ must equal V_2; hence

$$T = \frac{2V_2C}{I_0} = \frac{1}{f}, \qquad (11.7\text{–}1)$$

or equivalently,

$$\omega_0 = \frac{I_0\pi}{V_2C}. \qquad (11.7\text{–}2)$$

In the above analysis we assumed $I_0 > I_C$. If this were not the case, v_{o1} and v_{o2} would not drop to zero when the respective amplifiers were on, but rather would remain at some voltage which would probably be strongly dependent on the individual amplification factors of the components making up the amplifiers. In addition, this voltage which would probably be sensitive to temperature and power supply voltage, would determine the size of the negative step in v_{i1} and v_{i2}, and thus would strongly influence the period T. Consequently, in any rational design I_0 is chosen to be greater than I_C.

Now if the current drives in the circuit of Fig. 11.7-1 are permitted to be functions of time with the restriction

$$I(t) > I_C \qquad \text{for all } t, \qquad (11.7\text{–}3)$$

then the basic waveforms of Fig. 11.7-3 remain the same except that the v_i waveforms for $v_i < 0$ become frequency-modulated and take the form

$$v_{ij}(t) = -(V_2 - V_\delta) + \frac{1}{C}\int_{t_i}^{t} I_0\left[1 + \frac{I_1}{I_0}f(\theta)\right]d\theta, \qquad j = 1 \text{ or } 2, \qquad (11.7\text{–}4)$$

where $I(t) = I_1 f(t)$ and t_i is the time at which the circuit switched from its previous state. We may rewrite Eq. (11.7–4) in the form

$$v_{ij}(t) = -(V_2 - V_\delta) + \frac{I_0}{C}[\tau(t) - \tau(t_i)], \qquad (11.7\text{–}5)$$

vhere $\tau(t) = t + (I_1/I_0)\int^t f(\theta)\,d\theta$; hence $v_{ij}(t)$ vs. τ is a linear ramp with slope I_0/C vhich starts at $-(V_2 - V_\delta)$ and terminates at V_δ, at which point the circuit switches states. Clearly, then, the waveforms v_{i1}, v_{o1}, v_{i2}, and v_{o2} vs. $\tau(t)$ are identical with the orresponding waveforms vs. t for the case where $I(t) = I_0$. Consequently the output square wave is periodic in τ with period $T = 2V_2C/I_0$(or $\omega_0 = \pi I_0/V_2C$), which is he required form for an FM square wave with instantaneous frequency

$$\omega_i(t) = \frac{d}{dt}\omega_0\tau(t) = \omega_0\left[1 + \frac{I_1}{I_0}f(t)\right]. \qquad (11.7\text{–}6)$$

f I_1/I_0 is adjusted to equal $\Delta\omega/\omega_0$, then

$$\omega_i(t) = \omega_0 + \Delta\omega f(t). \qquad (11.7\text{–}7)$$

In the case of an FM square wave generator, it is even more essential to keep $(t) > I_C$. If this requirement is not met, the peak-to-peak amplitude of $v_o(t)$ will ary with time as $I(t)$ varies relative to I_C. In addition, since T is a function of both (t) and the peak-to-peak amplitude of v_o, the instantaneous frequency will not vary inearly with $I(t)$.

xample 11.7–1 Given two amplifiers with $V_2 = 5$ V, $V_\delta = 0$ V, and $I_C = 1$ mA, hoose values for $I(t) = I_0 + I_1 f(t)$ and C such that the circuit of Fig. 11.7–1 produces n FM square wave with $\omega_i(t) = 10^8 + 5 \times 10^7 f(t)$ rad/sec.

olution. Since $\Delta\omega/\omega_0 = I_1/I_0 = \frac{1}{2}$,

$$I(t)_{\min} = I_0 - I_1 = I_0/2.$$

hus we must choose $I_0 > 2$ mA to keep a constant peak-to-peak amplitude output quare wave. A choice of 3 mA for I_0 would be reasonable. Hence $I_1 = 1.5$ mA n addition, since $\omega_0 = 10^8$ rad/sec and since $\omega_0 = \pi I_0/V_2C$, $C = 19$ pF.

The problem now remains of implementing the amplifiers and current sources f Fig. 11.7–1 with practical circuits. One possible amplifier configuration employing iscrete transistors is shown in Fig. 11.7–4. For this circuit Q_1 provides the basic mplification, whereas Q_2, functioning as an emitter follower, provides the low utput impedance necessary to ensure that $v_o(t)$, which is loaded by the capacitor C

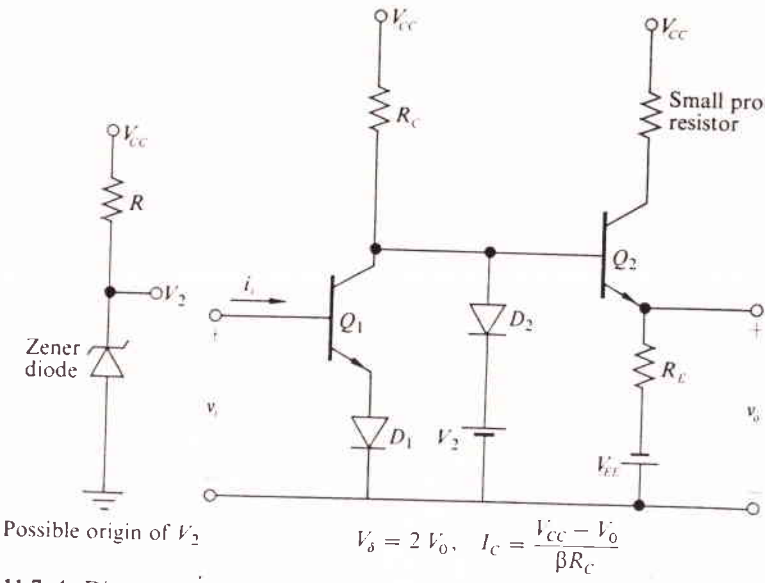


Fig. 11.7–4 Discrete component implementation of amplifiers in the circuit of Fig. 11.7–1.

rises rapidly when Q_1 is cut off. Specifically, the time constant controlling the rise time is closely approximated by

$$\tau = [R_C(1 - \alpha)\| R_E]C.$$

The diode-voltage-source combination (D_2-V_2) performs two functions. First, it catches the output voltage as it rises toward its steady-state value, thus further improving the rise time of the output square wave. Second, it establishes the peak output voltage at approximately V_2 ($V_2 < V_{CC}$), which is independent of supply voltage variations if V_2 is developed by a Zener diode. Hence $\omega_0 = \pi I_0/V_2C$ is less sensitive to power supply variations.

The diode D_1 provides a compensating junction voltage drop for that of Q_2 when Q_1 is saturated, thereby ensuring $v_o \approx 0$ when Q_1 is turned on. If D_1 were not included, v_o would vary by $V_2 + V_0$ instead of V_2 when Q_1 switches between saturation and cutoff; hence V_2 would be replaced by $V_2 + V_0$ in determining ω_0, and ω_0 would be significantly more dependent on temperature.

If for the circuit of Fig. 11.7–4 we assume that the base-emitter junctions of Q_1 and Q_2 and the diodes D_1 and D_2 can be modeled as ideal diodes in series with reverse-biasing batteries of $V_0 \approx 0.7$ V (for silicon transistors and diodes), then we observe that the v_i-i_i characteristic of the amplifier is identical to that of Fig. 11.7–2 with $V_\delta = 2V_0$. The forward transfer characteristic of the amplifier is also identical to that of Fig. 11.7–2 with $I_C = (V_{CC} - V_0)/\beta R_C$, where $\beta = \alpha/(1 - \alpha)$. Note that I_C is the value of the base current to Q_1 which causes Q_1 to saturate; that is, $v_{CE1} \approx 0$. Since little or no signal passes between the collector and the base of Q_1 (even without the emitter follower), the reverse transfer impedance is essentially zero while the output impedance, because of the emitter follower, is quite small.

To keep the output impedance small, the emitter follower must remain in its active region at all times. However, when the amplifier is operating within the astable multivibrator, and when Q_1 is saturated ($v_o \approx 0$), the current $I(t)$ must flow into the $v_o(t)$ terminal from the capacitor which is being charged. Thus to keep Q_2 active, the current through R_E must exceed $I(t)$ or

$$\frac{V_{EE}}{R_E} > |I(t)|_{\max}. \tag{11.7–8}$$

Equation (11.7–8) determines the maximum size of R_E. For example, if $|I(t)|_{\max} = 4.5$ mA (Example 11.7–1) and $V_{EE} = 9$ V, R_E must be chosen to be less than 2 kΩ. The resistor R_C is usually chosen somewhere between 200Ω and 5 kΩ; this is a compromise between fast switching times (small resistors are not affected as much as large resistors by stray capacitance) and power dissipation.

Figure 11.7–5 illustrates a complete FM square wave generator employing the amplifiers of Fig. 11.7–4. For this generator Q_5 and Q_6 act as the current sources $I(t)$. The V_3 and V_4 voltage sources provide the bias for Q_5 and Q_6, while D_5 (which is turned on by R) provides a compensating junction voltage for the base-emitter voltages of Q_5 and Q_6 so that some degree of thermal stability is obtained.

Again, for the circuit of Fig. 11.7–5, if the base-emitter junctions of Q_5 and Q_6 as well as D_5 are modeled by an ideal diode-battery combination, then

$$I(t) = \alpha\overbrace{\left(\frac{V_4 - V_3}{R_1} - \frac{V_3}{R_2}\right)}^{I_0} + \overbrace{\frac{\alpha V_1}{R_2}}^{I_1} f(t). \tag{11.7–9}$$

Since both V_3 and V_4 play a role in determining $I(t)$, both should be Zener controlled to keep $I(t)$ and in turn $\omega_i(t)$ independent of variations in supply voltage. In addition, the inequality

$$V_3 > V_\delta \tag{11.7–10}$$

must be satisfied to keep Q_5 and Q_6 from saturating when either Q_1 or Q_3 is on.

The resistors R_1 and R_2 are chosen to yield the desired values of I_0 and I_1 once V_1, V_3, and V_4 are selected. For example, if $V_1 = 3$ V, $V_3 = 2$ V, $V_4 = 10$ V, and $\alpha = 1$, while (from Example 11.7–1) $I_0 = 3$ mA and $I_1 = 1.5$ mA, then

$$R_2 = \frac{V_1}{I_1} = 2\text{ k}\Omega \qquad \text{and} \qquad R_1 = \frac{V_4 - V_3}{V_3/R_2 + I_0} = 2\text{ k}\Omega.$$

The circuit of Fig. 11.7–5 also illustrates a simple bandpass filter for extracting the fundamental FM wave from the FM square wave $v_{sq}(t)$. Specifically, if the filter BW $= 1/R_L C$ is sufficiently wide to pass the fundamental FM component of $v_{sq}(t)$ while attenuating all other FM components, then the filter output $v(t)$ is given by

$$v(t) = \frac{2V_2}{\pi}\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta + \theta_0\right], \tag{11.7–11}$$

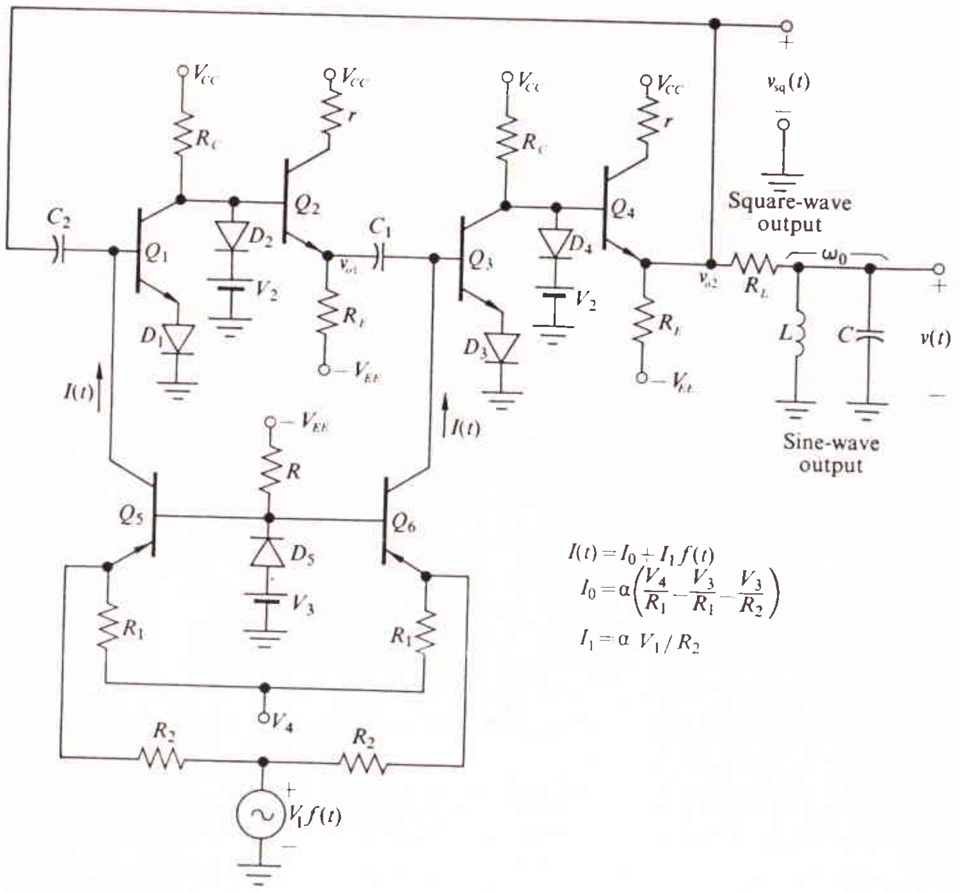


Fig. 11.7–5 Complete FM square-wave generator.

where

$$\omega_0 = \frac{\pi I_0}{V_2 C} = \frac{\pi}{V_2 C}\left(\frac{V_4 - V_3}{R_1} - \frac{V_3}{R_2}\right)$$

and

$$\Delta\omega = \frac{\pi I_1}{V_2 C} = \frac{\pi}{V_2 C}\frac{V_1}{R_2}.$$

(Note that the output filter shown has a transfer function of unity at ω_0.)

11.8 MISCELLANEOUS FREQUENCY MODULATORS—THE ARMSTRONG METHOD

In this section we shall consider several additional methods of producing FM signals. These methods include the Armstrong method and the method of varying the electrical time delay of a delay line. Neither of these methods is so versatile as the previously discussed methods, since they are quite restricted in the maximum size of both $\Delta\omega$ and ω_m relative to ω_0. The Armstrong method (named after its inventor E. H. Armstrong, who was a pioneer in the field of frequency modulation) has the advantage of carrier frequency stability, since its carrier may be obtained from a crystal oscillator. In addition, it is of interest historically, since it was one of the first methods of FM generation.

The variation of the electrical time delay of a delay line is of interest because this technique appears to be the method by which the frequency modulation of a laser beam can be most easily accomplished. Specifically, if a strong transverse electric field is placed across specially constructed transparent crystals, usually cadmium sulphide, the dielectric constant and, in turn, the velocity of wave propagation through the crystal vary with changes in the applied electric field. Since the crystal has a finite length d, variations in the propagation velocity v result in corresponding variations in the time taken by an electromagnetic wave, laser light in this case, to pass through the crystal; i.e. the time delay through the crystal is given by

$$t_0 = \frac{d}{v}. \tag{11.8–1}$$

Hence varying the applied field as a function of time produces corresponding variations in $t_0(t)$ which, as we shall see, are sufficient to produce a frequency-modulated signal.

The Armstrong method is a direct result of the fact that, for very small frequency deviations, an FM signal of the form

$$v(t) = A\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right], \tag{11.8–2}$$

which may be expanded as

$$v(t) = A\cos\omega_0 t\cos\Delta\omega\int^t f(\theta)\,d\theta - A\sin\omega_0 t\sin\Delta\omega\int^t f(\theta)\,d\theta, \tag{11.8–3}$$

can be closely approximated by

$$v(t) \approx A\cos\omega_0 t - A\,\Delta\omega\int^t f(\theta)\,d\theta\sin\omega_0 t. \tag{11.8–4}$$

Equation (11.8–4) follows from the fact that, for $\phi \leq 0.2$, $\cos\phi \approx 1$ and $\sin\phi \approx \phi$ within 2%. However, Eq. (11.8–4) may be realized by the block diagram of Fig. 11.8–1, which combines $A\cos\omega_0 t$ with the AM signal

$$A\,\Delta\omega\int^t f(\theta)\,d\theta\sin\omega_0 t.$$

Since ω_0 may be crystal controlled, the carrier frequency is quite stable.

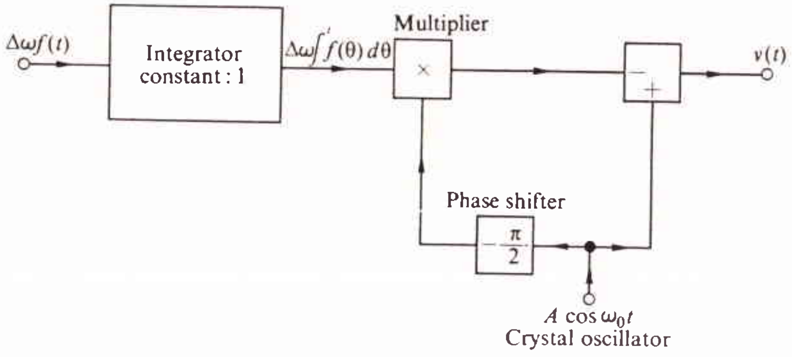


Fig. 11.8–1 Block diagram of Armstrong system for generating FM.

The only difficulty is that $\Delta\omega$ for the system of Fig. 11.8–1 of necessity must be quite small. Consider, for example, $f(t) = \cos\omega t$. For this case

$$\Delta\omega\int^t f(\theta)\,d\theta = \frac{\Delta\omega}{\omega}\sin\omega t,$$

which has a peak value of $\Delta\omega/\omega$. For the Armstrong modulator to function correctly in this case, $\Delta\omega/\omega$ must be less than or equal to 0.2. For the more general case in which $f(t)$ is band-limited between ω_L and ω_m, $\Delta\omega\int^t f(\theta)\,d\theta$ has a peak value ranging from a maximum of $\Delta\omega/\omega_L$ [when all the energy of $f(t)$ is concentrated at ω_L] to a minimum of $\Delta\omega/\omega_m$ [when all the energy of $f(t)$ is concentrated at ω_m]. However, since the peak value of $\Delta\omega\int^t f(\theta)\,d\theta$ must be less than 0.2 for all possible frequency distributions of $f(t)$, $\Delta\omega$ must be sufficiently small that

$$\frac{\Delta\omega}{\omega_L} \leq 0.2. \tag{11.8–5}$$

In practice, Eq. (11.8–5) is usually relaxed to

$$\frac{\Delta\omega}{\omega_L} < 0.5, \tag{11.8–6}$$

since very rarely is all the energy of $f(t)$ concentrated at ω_L. Even Eq. (11.8–6) requires extremely small values of $\Delta\omega$. If $f(t)$ is band-limited between 100 Hz and 10 kHz, $\Delta\omega$ must be less than 2π(50 Hz). Since $\Delta\omega$ is directly determined by the value of ω_L, the minimum value of ω_L should be no smaller than necessary when using the Armstrong modulator.

To overcome the problem of the small frequency deviation appearing at the output of the Armstrong modulator, we may place the output signal through a sequence of frequency doublers and triplers† (of the form discussed in Section 1.4).

---

† Higher-order multiplication is usually not attempted because of the difficulty in extracting the higher harmonics at the multiplier output with conventional tuned circuits.

reby increasing the frequency deviation, as was pointed out in Section 11.2. For mple, if the Armstrong modulator produces an FM signal with $\omega_0 = 2\pi(100\text{ kHz})$ l $\Delta\omega = 2\pi(50\text{ Hz})$, by placing ten frequency doublers after the modulator we ain at the output of the last doubler a frequency deviation of $2\pi(50\text{ Hz}) \times 2^{10} =$ 51.2 kHz). In addition, we obtain a carrier frequency of $2\pi(100\text{ kHz}) \times 2^{10} =$ 102.4 MHz). This carrier may be heterodyned to any desired frequency by any the techniques of Chapter 7 with no effect on the frequency deviation. Another stal-controlled oscillator should be employed in the mixer to maintain the carrier quency stability.

Since all the blocks of the diagram of Fig. 11.8–1 can be physically implemented circuits discussed elsewhere in this book, we leave the physical circuit construction an exercise for the reader.

riable Time-Delay Modulator

the delay experienced by a constant-amplitude sinusoid $A \cos \omega_0 t$ passing through network is a function of time, i.e., if $t_0 = t_0(t)$, then the signal appearing at the twork output has the form

$$v(t) = KA \cos[\omega_0 t - \omega_0 t_0(t)], \tag{11.8–7}$$

ere K is the attenuation introduced by the network. Now if $t_0(t)$ can somehow be used to vary in proportion to $\int^t f(\theta)\, d\theta$, i.e., if

$$t_0(t) = \frac{\Delta\omega}{\omega_0} \int^t f(\theta)\, d\theta, \tag{11.8–8}$$

en an FM signal of the desired form is produced at the network output. In the case the frequency modulation of a laser beam, the field across the crystal through hich the beam is passed must be varied in proportion to $\int^t f(\theta)\, d\theta$ to produce a ne delay of the form given by Eq. (11.8–8).

.9 FREQUENCY STABILIZATION OF FREQUENCY MODULATORS

many applications the carrier frequency at the output of an FM generator is quired to be extremely stable. For example, the Federal Communications Comission requires that commercial FM stations (88 MHz to 108 MHz) maintain a rrier frequency stability of ±2 kHz about their assigned station frequency. This orresponds to maximum carrier variation of two parts in 10^5 for a station at 00 MHz. Such stability is usually not directly obtainable with any of the FM enerators previously discussed, with the possible exception of the crystal controlled rmstrong generator.

This stability may be obtained by comparing the output signal of an FM generator ith a crystal oscillator in such a fashion that an error signal results if the FM carrier not at its specified frequency. This error signal is then amplified and fed back to he FM generator input with the correct phase to return the FM carrier to its required equency.

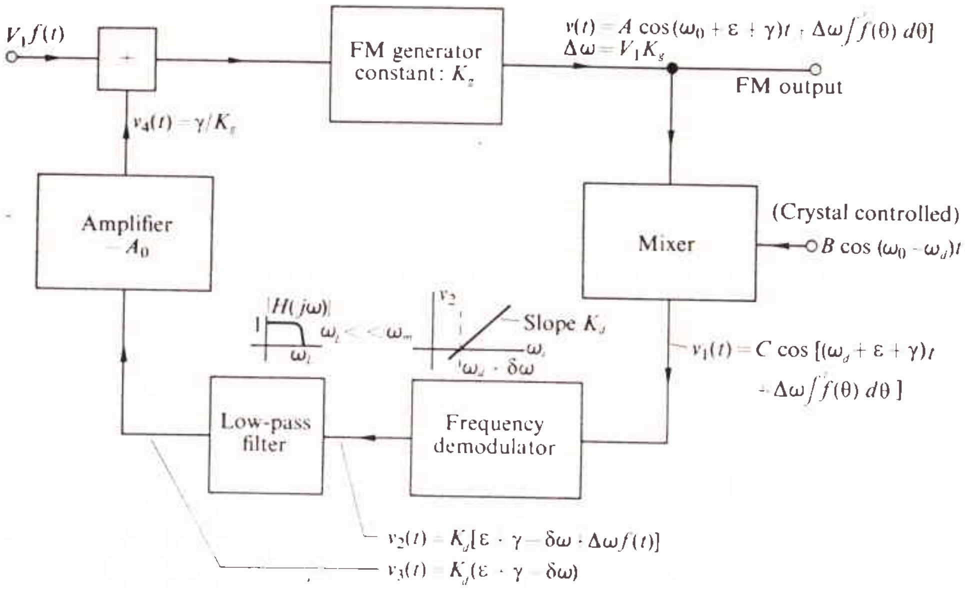

$V_1 f(t)$
+
FM generator constant: K_g
$v(t) = A\cos[(\omega_0 + \varepsilon + \gamma)t + \Delta\omega \int f(\theta)\, d\theta]$
$\Delta\omega = V_1 K_g$
FM output
$v_4(t) = \gamma / K_g$
Amplifier
A_0
Mixer
(Crystal controlled)
$B\cos(\omega_0 - \omega_d)t$
$|H(j\omega)|$
1
$\omega_l < < \omega_m$
ω_l
v_2
Slope K_d
ω_i
$\omega_d + \delta\omega$
$v_1(t) = C\cos[(\omega_d + \varepsilon + \gamma)t + \Delta\omega \int f(\theta)\, d\theta]$
Low-pass filter
Frequency demodulator
$v_2(t) = K_d[\varepsilon + \gamma - \delta\omega + \Delta\omega f(t)]$
$v_3(t) = K_d(\varepsilon + \gamma - \delta\omega)$

Fig. 11.9–1 Block diagram of frequency stabilization system.

A widely used frequency stabilization feedback system is shown in block diagram form in Fig. 11.9–1. In this figure the output of the FM generator is assumed to be of the form

$$v(t) = A \cos\left[\omega_0 t + \varepsilon t + \gamma t + \Delta\omega \int^t f(\theta)\, d\theta\right], \tag{11.9–1}$$

where ω_0 is the desired carrier frequency, γ is the frequency shift due to the feedback loop, and ε is the frequency difference between the actual carrier frequency with no feedback and the desired carrier frequency. The FM generator output is mixed with a highly stable crystal-controlled oscillator and the difference term is extracted. If the crystal-controlled oscillator operates at the frequency $\omega_0 - \omega_d$, then the mixer output has the form

$$v_1(t) = C \cos\left[(\omega_d + \varepsilon + \gamma)t + \Delta\omega \int^t f(\theta)\, d\theta\right]. \tag{11.9–2}$$

The mixer output is then placed through a frequency demodulator (to be discussed in Chapter 12) which has the transfer characteristic between instantaneous frequency and output voltage shown in Fig. 11.9–1, with the result that the demodulator output is given by

$$v_2(t) = K_d[\varepsilon + \gamma - \delta\omega + \Delta\omega f(t)]. \tag{11.9–3}$$

n reality, the center frequency of the discriminator should be ω_d. The term $\delta\omega$ epresents the error in obtaining the desired frequency. This term combines the accuracies of the demodulator and the crystal oscillator.)

The discriminator output is placed through a low-pass filter which is sufficiently arrow to remove the $f(t)$-term from $v_2(t)$. The filter output is then amplified by A_0 and applied to the input of the FM generator. To produce a term γt in the hase of the FM generator output, the filter output must have the value γ/K_g as well s $-A_0K_d[\varepsilon + \gamma - \delta\omega]$; thus

$$\gamma = -A_0K_gK_d[\varepsilon + \gamma - \delta\omega], \tag{11.9–4}$$

here K_g is the generator constant which relates input voltage to frequency deviation.

Equation (11.9–4) may be rearranged in the form

$$\gamma = \frac{-A_0K_gK_d(\varepsilon - \delta\omega)}{1 + A_0K_gK_d} = \frac{-\varepsilon + \delta\omega}{1 + 1/A_0K_gK_d}. \tag{11.9–5}$$

rom Eq. (11.9–5) it is apparent that, as $A_0K_gK_d$ approaches infinity, γ approaches $\varepsilon + \delta\omega$ and the resultant net carrier frequency approaches $\omega_0 + \delta\omega$. Thus for a igh-gain feedback system the final stability is a function of only the crystal oscillator tability and the demodulator stability.

Once one has a high-gain (nonoscillating) system, one has the problem of spreading the allowable variations between the crystal oscillator and the frequency demoduator. If the reference frequency is very close to the desired output frequency, then ω_d will be much lower than ω_0 and a 1% detuning of the demodulator will cause a nuch smaller percentage change in ω_0.

As we shall see, frequency demodulators can be constructed of quartz crystal uning elements; hence, with proper temperature control, stabilities of several parts er million are available both from the initial crystal oscillator and from the discriminator. As an example, consider a case where $A_0K_gK_d = 100$ and where an FM enerator at 20 MHz is used to generate a final carrier at 60 MHz. The stabilizing rystal oscillator operates at 5.9 MHz, and this frequency is multiplied by a factor f 10 to obtain the reference frequency. (In a practical circuit, multiplication by × 3 = 9 or 2 × 2 × 3 = 12 would be more reasonable.) The frequency demodulator perates at 1 MHz.

From a modified version of Eq. (11.9–4),

$$\gamma = -100(3\varepsilon + 3\gamma - \delta_{\text{disc}}\omega - 10\,\delta_{\text{osc}}\omega)$$

r

$$\gamma = \frac{-300}{301}\varepsilon + \frac{100}{301}\delta_{\text{disc}}\omega + \frac{1000}{301}\delta_{\text{osc}}\omega$$

lence, if the original 20 MHz oscillator has a maximum drift of 1% or $\varepsilon = 2$ 200 kHz), then the final 60 MHz signal has the following "drift" terms:

$$\frac{600}{301}\,\text{kH} + \frac{300}{301}\delta_{\text{disc}}f + \frac{3000}{301}\delta_{\text{osc}}f.$$

If the discriminator is accurate and stable to within $\frac{1}{4}$% and the 5.9 MHz oscillator is stable to within 15 parts per million, then the final worst-case output variation from 60 MHz should not exceed approximately

$$2\,\text{kHz} + 2.5\,\text{kHz} + 900\,\text{Hz} = 5.4\,\text{kHz},$$

or less than one part in 10^4. In this case a better crystal oscillator would not be of much help until one had raised the feedback gain and improved the discriminator stability.

PROBLEMS

11.1 For the sinusoidally modulated FM signal

$$v(t) = V_1 \cos(\omega_0 t + \beta \sin \omega_m t),$$

a) show that the rms value of $v(t)$ is given by

$$V_{\text{rms}} = \sqrt{\frac{V_1^2}{2}\left[J_0^2(\beta) + 2\sum_{n=1}^{\infty} J_n^2(\beta)\right]}$$

$$= \frac{V_1}{\sqrt{2}}$$

(*Hint*: $\sum_{n=-\infty}^{\infty} J_n^2(\beta) = 1$) and

b) determine the amplitude of the significant sidebands for the case where $\beta = 5$ and $V_1 = 5$ V.

11.2 Determine the bandwidth occupied by the phase-modulated signal

$$v(t) = A\cos[\omega_0 t + \Delta\phi f(t)]$$

for the case where $\Delta\phi = 5$ rad and $f(t) = \sin 10^5 t$. Retain only frequency components which are greater than 1% of the unmodulated carrier.

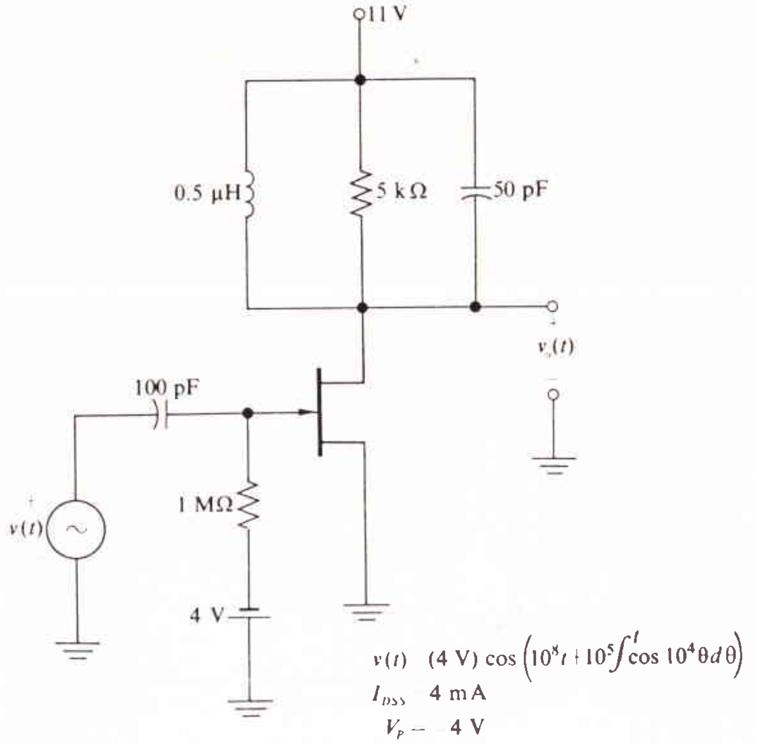
11.3 A 10 V peak-to-peak FM square wave is generated with a carrier frequency of 14 MHz and modulated with a baseband signal which is band-limited to 5 MHz. The FM square wave is then placed through a rectangular bandpass filter centered at 70 MHz in order to obtain a 70 MHz FM test signal. The filter BW passes all sidebands greater than 1% of the unmodulated carrier at 70 MHz. Determine the maximum frequency deviation which may be applied to the FM square wave without introducing excessive distortion into the 70 MHz test signal.

11.4 For the circuit shown in Fig. 11.P–1, determine an expression for $v_o(t)$. Assume that the FET has a square-law characteristic in its saturation region.

11.5 The FM signal

$$v(t) = A\cos\left(\omega_0 t + \frac{\Delta\omega}{\omega_m}\sin\omega_m t\right)$$

is passed through a high-Q single-tuned filter with bandwidth BW $= 2x$ centred at ω_0.



$v(t)$ (4 V) $\cos\left(10^8 t + 10^5 \int^t \cos 10^4 \theta \, d\theta\right)$
I_{DSS} 4 mA
V_P = −4 V

Figure 11.P–1

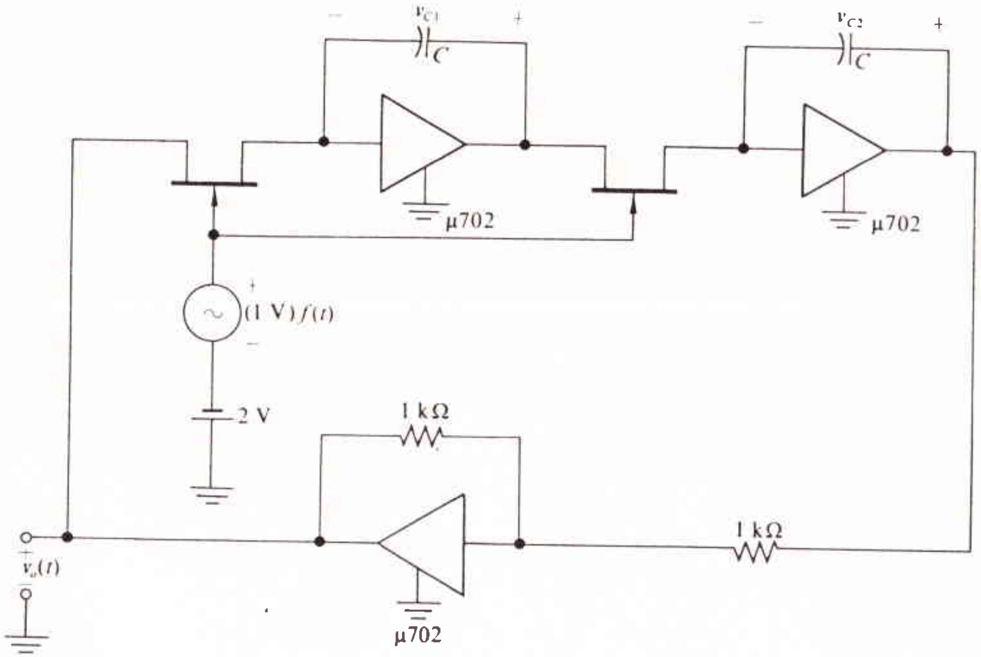
Assuming the "quasi-static" approximation to be valid, show that the instantaneous frequency at the filter output is given by

$$\omega_i(t) = \omega_0 + \Delta\omega \cos \omega_m t - \frac{d}{dt}\left(\tan^{-1} \frac{\Delta\omega \cos \omega_m t}{\alpha}\right).$$

For the case where $\Delta\omega/\omega_m = \beta = 5$ and $\Delta\omega/\alpha = \frac{1}{4}$ (a set of parameters for which the quasi-static approximation is valid), determine the percentage of third-harmonic distortion (relative to $\Delta\omega$) introduced into the instantaneous frequency by the nonlinear phase characteristic of the filter. (*Hint:* Expand $\tan^{-1}\theta$ in its MacLauren series and keep the first two terms.)

11.6 Verify that $v(t) = C \cos\left[\int^t \omega_i(\theta)\, d\theta + \theta_0\right]$ is indeed a solution of Eq. (11.4–4).

11.7 For the circuit shown in Fig. 11.P–2, the $t = 0$ value of v_{C1} is 100 mV while the $t = 0$ value of v_{C2} is 0. Determine an expression for $v_o(t)$ for $t \geq 0$ (cf. Section 8.3). (In practice, the circuit of Fig. 11.P–2 has a sufficient number of leakage paths to be self-starting. In addition the oscillation amplitude is limited by the nonlinearity of the drain-source resistance the FET at about 200 mV.)



$I_{DSS} = 4$ mA, $V_P = -4$ V, $C = 100$ pF

Operational amplifier gain ≈ 3000

Figure 11.P–2

11.8 For the circuit shown in Fig. 11.P–3, Q_1 has a collector-base capacitance given by $C(v_{CB}) = Av_{CB}^{-1/2}$, where

$$A = 15.7 \text{ pF} \sqrt{\text{volts}}.$$

a) Show that the instantaneous frequency of the oscillator is given by

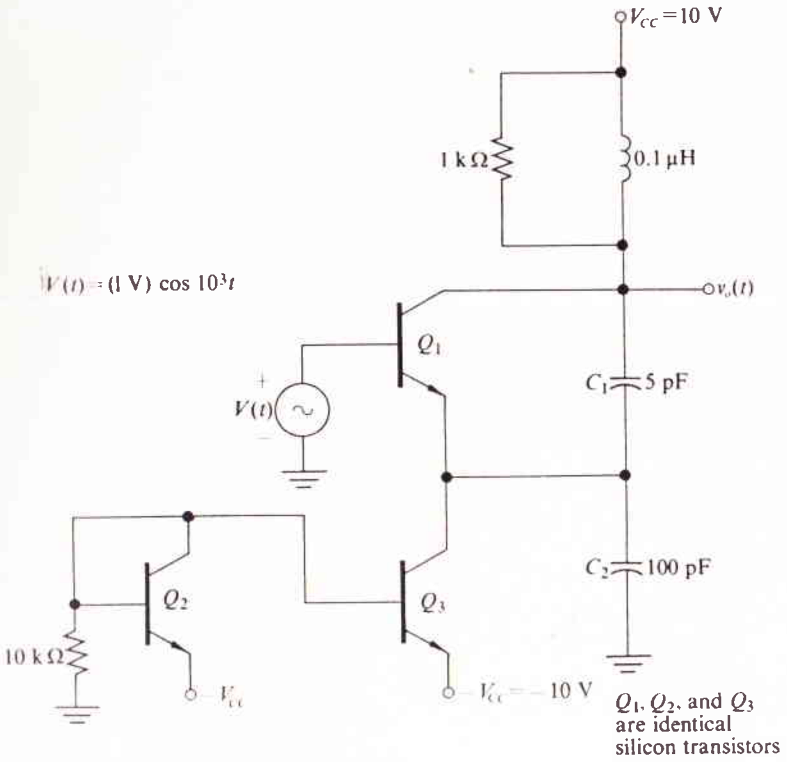
$$\omega_i(t) = \frac{1}{\sqrt{L[C_0 + C(t)]}},$$

where

$$C_0 = \frac{C_1 C_2}{C_1 C_2} + C_0', \qquad C_0' = A V_{CC}^{-1/2}, \qquad \text{and} \qquad C(t) = \frac{A}{2} V_{CC}^{-3/2} V(t).$$

b) Determine a numerical expression for $v_o(t)$. (In a practical FM oscillator with small values of C_1 and C_2, most of C_2 would be supplied by the emitter-base capacitance of the transistor and thus the physical capacitor C_2 would be omitted.)

11.9 Verify that Eq. (11.4–14) is the solution to Eq. (11.4–13).



Q_1, Q_2, and Q_3 are identical silicon transistors

Figure 11.P–3

11.10 Determine expressions for $v_i(t)$ and $v_o(t)$ in the circuit of Fig. 11.5–2 with the following parameter values:

$$V_{EE} = V_{CC} = 15\text{ V}, \quad I_{k0} = I_{k1} = 1\text{ mA}, \quad I_k = 2\text{ mA}, \quad C_0 = 0,$$

$$f(t) = \cos 5 \times 10^2 t, \quad R_{\text{loss}} = R_L = 5\text{ k}\Omega, \quad r = \tfrac{1}{10}\,\Omega,$$

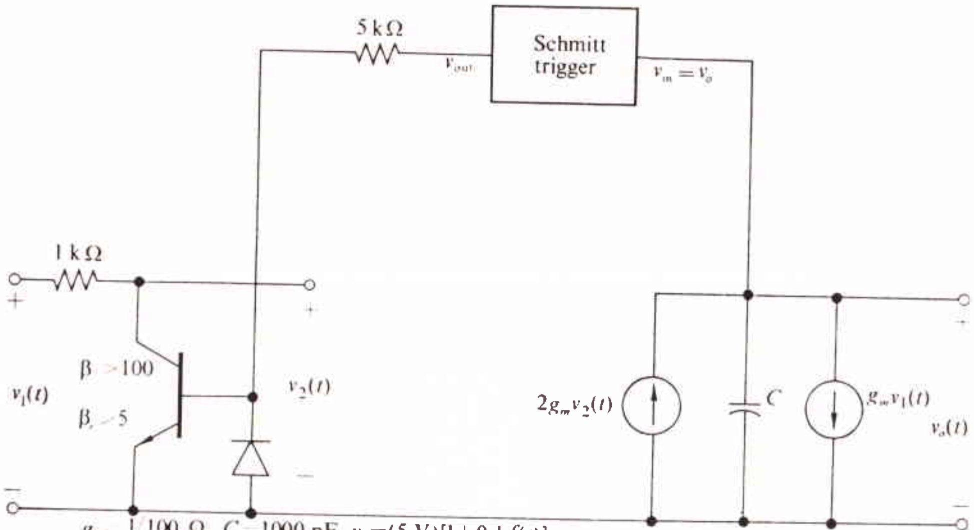
$$C = C_L = 1000\text{ pF}, \quad L = L_L = 10\ \mu\text{H}, \quad \text{and} \quad M = 0.5\ \mu\text{H}.$$

All transistors are silicon and identical.

11.11 For the triangular-wave FM generator shown in Fig. 11.P–4, determine the instantaneous frequency. The Schmitt trigger in the circuit has the following characteristics:

1) When $v_o(t)$ increases upward through +5 V, the Schmitt trigger output switches from −10 V to +10 V.
2) When $v_o(t)$ decreases downward through −5 V, the Schmitt trigger switches from +10 V to −10 V.

11.12 Design an eight-diode wave-shaping network similar to the one shown in Fig. 11.6–9(b) which converts the output of the circuit of Fig. 11.P–4 into a 4 V peak-to-peak sinusoidal FM wave. The time intervals between break points on the output waveform should be equal.



$g_m = 1/100\ \Omega$, $C = 1000$ pF, $v_1 = (5\text{ V})[1 + 0.1 f(t)]$

Figure 11.P–4

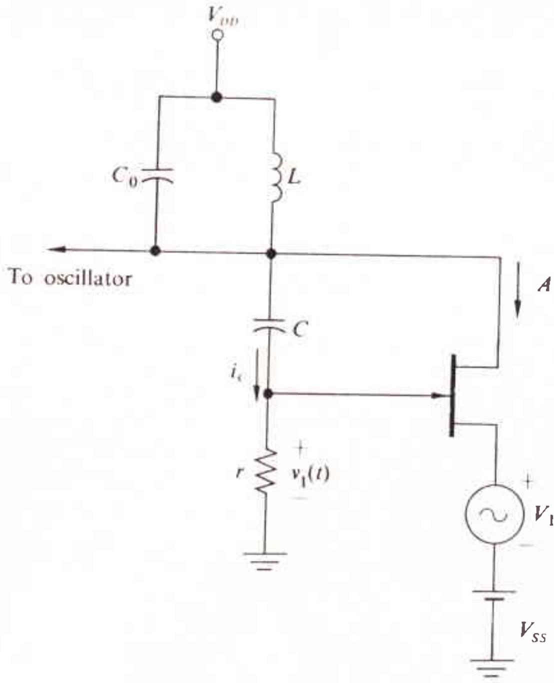


Figure 11.P–5

11.13 Show that the FET circuit in Fig. 11.P–5 may be modeled as a controlled current source $A(t)i_c$ where

$$A(t) = g_m(t)r = \overbrace{-r\frac{2I_{DSS}}{V_P}\left(1 + \frac{V_{SS}}{V_P}\right)}^{A_0} - \overbrace{r\frac{2I_{DSS}}{V_P}\left(\frac{V_1 f(t)}{V_P}\right)}^{A_1 f(t)},$$

provided that

a) $r \ll 1/\omega_0 C$,
b) $v_1(t) \ll |V_P|$,
c) the FET operates in its square-law region, and
d) the low-frequency components of i_D are shunted to ground by L.

If $C_0 = C = 500$ pF, $L = 10\ \mu$H, $r = 1\ \Omega$, $V_P = -4$ V, $V_{SS} = 2$ V, $V_1 = 1$ V, $V_{DD} = 10$ V, $I_{DSS} = 4$ mA, and the oscillator limits the output waveform to 4 V peak-to-peak, find an expression for the FM voltage across the tuned circuit for the case where $f(t) = \cos 10^3 t$.

11.14 Design an FET oscillator which can be used in conjunction with the circuit shown in Fig. 11.P–5 which has the numerical values given in Problem 11.13.

11.15 In the circuit shown in Fig. 11.7–5, $V_{CC} = V_{EE} = V_4 = 12$ V, $V_2 = 5$ V, $V_1 = 1.5$ V, $V_3 = 3$ V, $R_1 = 4.5$ kΩ, $R_2 = 3$ kΩ, $R_E = 4$ kΩ, $R_C = 2$ kΩ, $r = 47\ \Omega$, $C_1 = C_2 = 1000$ pF, $\beta > 100$, and $f(t) = \cos 10^4 t$. Sketch $v_{sq}(t)$ and find an expression for the instantaneous frequency.

CHAPTER 12

FM DEMODULATORS

In broad outline an FM receiver is similar to an AM receiver. Both are normally superheterodyne types with RF amplifiers, RF filtering, mixers, and IF amplification and filtering. In the AM receiver the desired output information is carried in the carrier envelope fluctuations; hence envelope linearity must be maintained not only in the front end but throughout the receiver. In the FM receiver the desired information is contained in the variations in the instantaneous frequency; hence, in order to demodulate the FM signal, one must convert these frequency variations into baseband amplitude variations.

Unfortunately, most FM demodulators are at least to some extent also capable of demodulating AM signals. Hence if, in the course of passage through the transmission channel or the front end of the receiver, the envelope of the FM signal is corrupted by noise, fading or filter-induced FM-to-AM conversion, then such envelope corruptions will be detected and will appear at the output as baseband noise or distortion.

As we shall see, some FM detectors are more susceptible than others to this type of distortion. However almost all FM receivers of reasonable quality include an envelope amplitude variation removal circuit, commonly called a limiter, between the last IF amplifier and the demodulator. In some practical circuits, the ratio detector for example, the limiting and the demodulation functions become somewhat entangled. For the sake of initial understanding we shall first study demodulators for which these operations are completely separate and independent.

After a discussion of limiters, this chapter outlines the theoretical limitations on various types of FM demodulators. It then analyzes *all* the common FM demodulation circuits as well as several useful but not yet common circuits. The chapter ends with a section on the more complex detectors that are used in "threshold extension" FM receivers.

12.1 LIMITERS

The limiter circuit is ideally a zero-memory nonlinear circuit that produces a constant-amplitude, constant-waveshape type of output for all its permissible input levels or waveshape types. Figure 12.1–1 shows an ideal limiter characteristic; Fig. 12.1–2 shows two practical characteristics.

Any noise-corrupted or amplitude-modulated FM signal applied to the circuit of Fig. 12.1–1 will be converted into a constant-amplitude FM square wave. This

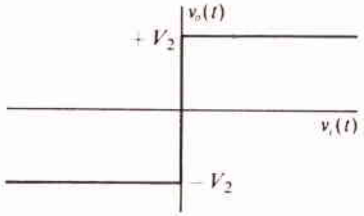
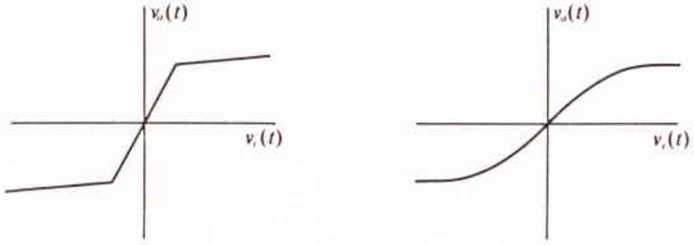


Fig. 12.1–1 Ideal limiter characteristic.



(a) (b)

Fig. 12.1–2 "Practical" limiter charactistics.

square wave may be applied directly to certain types of FM detectors. In other cases it may be filtered—the filter must remove the carrier's third and higher odd harmonics and their sidebands without disturbing the sidebands around the carrier—to produce a constant-amplitude sinusoidal FM signal. In actual circuits, such filters at the outputs of limiters tend to be designed primarily to satisfy the requirements of the discriminator rather than those of the limiter. Normally their bandwidths are just large enough to prevent signal-induced AM which causes distortion at the demodulator output.

Figure 12.1–3 shows a circuit that might be modeled by the characteristic of Fig. 12.1–2(a). To accomplish such a modeling, one replaces the real diodes by a combination of an ideal diode, a battery, and a series resistor. The two output voltage breakpoints occur when

$$v_o(t) = \pm\frac{v_i(t)R_0}{R_i + R_0} = \pm V_0,$$

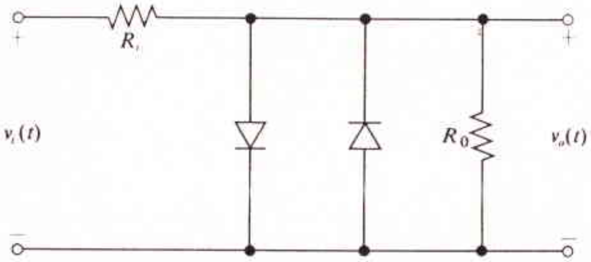


Fig. 12.1–3 Diode limiter.

where V_0 is the diode turn-on voltage. The midsection slope is $R_0/(R_0 + R_i)$, while the outside section slopes includes r_{diode} in parallel with R_0. It is obvious that, as r_{diode} approaches zero and as the input driving amplitude becomes very large with respect to V_0, the output waveshape approaches a square wave. Of course, even if r_{diode} is zero, there will be no limiting at all if the input amplitude is less than V_0. Whereas the rise and fall times of the output signal from the characteristic of Fig. 12.1–1 are always independent of the driving signal, the output slopes for the characteristic of Fig. 12.1–2(a) are always a function of the drive.

Though the output signal from Fig. 12.1–2 is not a square wave, its fundamental component can be calculated directly through the use of the sine-wave tip characteristics of Chapter 4. Thus the quasi-static output of such a limiter followed by a bandpass filter is known exactly.

With the symmetrical characteristics of Fig. 12.1–2 go the tremendous advantages of having no dc or even harmonics. However, various single-ended characteristics can also be used as limiters. For example, the $I_1(x)/I_0(x)$ curve of Fig. 4.5–5 indicates the input–filtered-output relationship for a constant-current bias-driven exponential junction. Such limiters are restricted in their modulating frequency capabilities; as was pointed out in Section 5.6, the input coupling capacitor requires that ω_0 be of the order of $10^4\omega_m$ if proper limiting action is to occur.

The present-day practical circuit that comes closest to approximating the ideal characteristic is a differential-pair limiter preceded by one or more stages of differential-pair drivers. Figure 4.6–2 indicates the basic characteristic, while Fig. 4.6–4 indicates the output of a bandpass filter following a differential-pair limiter. As indicated in Table 4.6–1, an input voltage swing of ± 78 mV swings the output current in either transistor from 5% to 95% (or vice versa) of its peak value. Thus if the limiter stage is preceded by another symmetrical stage with a voltage gain of 78, then a ± 1 mV input voltage swing will produce this nearly complete output swing. Such a characteristic does indeed approach the ideal curve of Fig. 12.1–1 quite closely.

Commercial integrated-circuit differential-pair limiters with three cascaded pairs are readily available. These units have voltage gains of 70 dB or more (greater than 3000 times) into a 1 kΩ load. Figure 12.1–4 illustrates a simplified version of such a limiter. This unit has a level-shifting emitter-follower output to allow it to supply a "square-wave" output which is centered about 0 V. If the circuit were used in conjunction with a bandpass filter, the filter would replace R_2 in the collector of Q_4.

If a sinusoidal FM signal is required at the limiter output, a bandpass filter centered at the carrier frequency must follow the nonlinear limiter characteristic. This filter must have sufficient bandwidth (usually BW > 8 $\Delta\omega$) so that its amplitude characteristic remains sufficiently flat in the vicinity of ω_0 to prevent deviation-induced amplitude variations (cf. Eq. 11.3–2). On the other hand, the filter bandwidth must be sufficiently narrow so that the higher harmonic components of the nonlinear limiter output do not enter its passband. Such conflicting requirements are possible only if D and β for a given FM signal determine a point well below the appropriate curve in Fig. 11.2–3. Problems 12.4 and 12.5 consider this point in more detail.

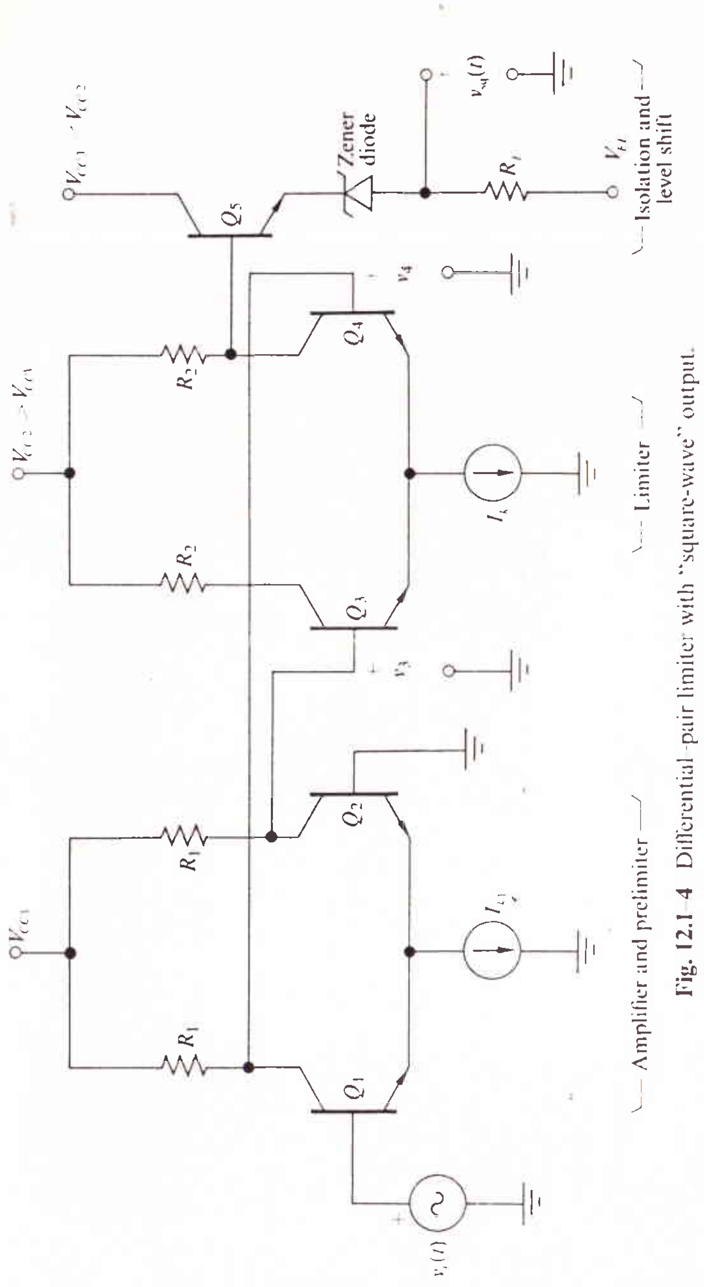


Fig. 12.1-4 Differential-pair limiter with "square-wave" output.

Dynamic Limiting

When a peak envelope detector is placed across a tuned circuit driven by an amplitude-modulated current as shown in Fig. 12.1–5, the modulation of the tuned-circuit voltage depends not only on the bandwidth of the parallel *RLC* circuit, but also on the R_0C_0 time constant of the envelope detector circuit. In particular, as was shown in Example (10.3–2), if the envelope detector time constant becomes sufficiently large, all amplitude modulation is stripped from the tuned-circuit voltage, provided that failure-to-follow distortion does not occur. In addition, if $i_i(t)$ is an amplitude-modulated FM signal, the amplitude variations may be removed from the tuned-circuit voltage while its instantaneous frequency remains identical with the instantaneous frequency of $i_i(t)$.

As we shall now demonstrate, even if $i_i(t)$ is an amplitude-modulated FM signal, the envelope of the tuned-circuit voltage $v(t)$ and the point at which diagonal clipping occurs may be determined from the equivalent circuit for the envelope detector developed in Section 10.3. At the same time we shall show that the instantaneous phase of $i_i(t)$ and $v(t)$ are identical, provided that the bandwidth of the unloaded tuned circuit is sufficiently large so that the circuit appears resistive (of value *R*) over the band of frequencies occupied by the FM signal.

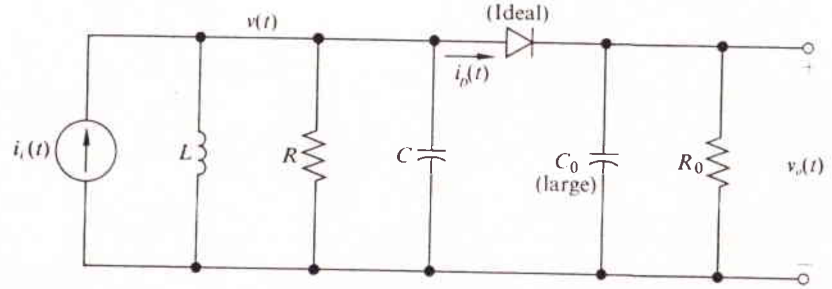
To demonstrate that the results of Section 10.3 remain valid when $i_i(t)$ is an amplitude-modulated FM signal of the form

$$i_i(t) = b(t)\cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\,d\theta\right]$$

$$= b(t)\cos \omega_0 \tau(t), \qquad (12.1\text{–}1)$$

we retrace the analysis of Section 10.3. We first assume that Q_T for the parallel *RLC* circuit is sufficiently high so that $v(t)$ contains frequency components only in the vicinity of ω_0. Therefore, $v(t)$ may be written in the general form

$$v(t) = g(t)\cos \omega_0 \tau_1(t).$$



$i_i(t) = b(t)\cos[\omega_0 \tau(t)]$

Fig. 12.1–5 Dynamic limiter.

Now if the R_0C_0 time constant is sufficiently long so that $v_o(t)$ contains little or no ripple and if no failure-to-follow distortion occurs, then $v_o(t)$ must equal $g(t)$. Also, the diode current must consist of a train of narrow pulses occurring at the peaks of $v(t)$; thus $i_D(t)$ may be expanded in a Fourier-like series of the form

$$i_D(t) = i_{D0}(t) + 2i_{D0}(t)\cos\omega_0\tau_1(t) + \cdots. \qquad (12.1\text{–}2)$$

Since $i_{D0}(t)$ flowing through an R_0C_0 circuit must develop $v_0(t) = g(t)$, $g(t)$ may be expressed as

$$g(t) = i_{D0}(t) * z_0(t). \qquad (12.1\text{–}3)$$

where $z_0(t)$ is the impulse response of the R_0C_0 network. In addition, the equating of the fundamental FM signal components flowing into the tuned circuit to zero results in

$$b(t)\cos\omega_0\tau = \left[\frac{g(t)}{R} + 2i_{D0}(t)\right]\cos\omega_0\tau_1(t). \qquad (12.1\text{–}4)$$

[Here it is assumed that the tuned circuit appears resistive to $v(t)$.] For Eq. (12.1–4) to be satisfied, $\tau_1(t)$ must be equal $\tau(t)$ and also

$$\frac{b(t)}{2} - \frac{g(t)}{2R} = i_{D0}(t). \qquad (12.1\text{–}5)$$

As in Section 10.3 the simultaneous equations, Eqs. (12.1–3) and (12.1–5), may be related to the equivalent circuit shown in Fig. 12.1–6. We note that this circuit is identical with the corresponding circuit of Fig. 10.3–5 [except that $C = 0$, which is a direct result of assuming the tuned circuit to be resistive to $v(t)$]; therefore, the envelope $g(t)$ and the point where failure-to-follow distortion occurs may be determined in exactly the same fashion as they were in Chapter 10.

It is now clear from Fig. 12.1–6 that, if C_0 is chosen sufficiently large so that it appears as an ac short circuit over the band of frequencies occupied by the ac component of $b(t)$, then $g(t)$ rises slowly to the dc level

$$g(t) = \overline{b(t)}\left(R \,\|\, \frac{R_0}{2}\right), \qquad (12.1\text{–}6)$$

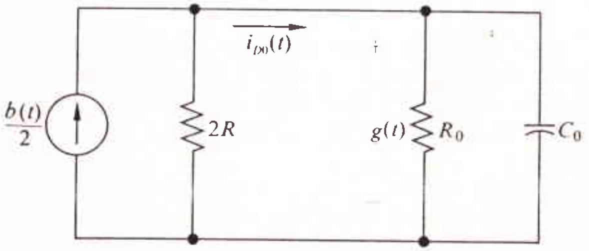


Fig. 12.1–6 Equivalent circuit for determining $g(t)$ and $i_{D0}(t)$.

where $\overline{b(t)}$ is the average envelope level of $i_i(t)$. Consequently, since $\tau_1(t) = \tau(t)$, $v(t)$ is a limited FM signal given by

$$v(t) = \underbrace{\overline{b(t)}\left(R \,\|\, \frac{R_0}{2}\right)}_{A}\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right]. \qquad (12.1\text{–}7)$$

Although the envelope of $v(t)$ is insensitive to rapid variations in $b(t)$, it does respond to slow changes in the average value of $b(t)$; hence the circuit limits only dynamically.

The limiting action of the circuit of Fig. 12.1–5 assumes that failure-to-follow distortion does not occur. It was determined in Eq. (10.3–4) that the condition which ensured no failure-to-follow distortion for the dynamic limiter was given by

$$m \le \frac{2R}{R_0 + 2R}.$$

Consequently, if large envelope variations in $b(t)$ are expected, R_0 must be kept small in comparison with R, at the expense of a greatly reduced output amplitude, to prevent failure-to-follow distortion.

12.2 FREQUENCY DEMODULATION TECHNIQUES

In this section we shall explore some of the methods by which a signal proportional to $f(t)$ can be obtained from a limited FM signal having either of the forms

$$v(t) = A\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right] = A\cos\omega_0\tau(t) \qquad (12.2\text{–}1\text{a})$$

or

$$v_{sq}(t) = BF\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right] = BF[\omega_0\tau(t)], \qquad (12.2\text{–}1\text{b})$$

where $\tau(t) = t + (\Delta\omega/\omega_0)\int^t f(\theta)\,d\theta$ and $F(\theta)$ is the periodic square wave shown in Fig. 12.2–1. The voltage $v(t)$ represents the output of a limiter followed by a bandpass filter to extract the fundamental FM signal, whereas $v_{sq}(t)$ represents the direct output of an ideal limiter.

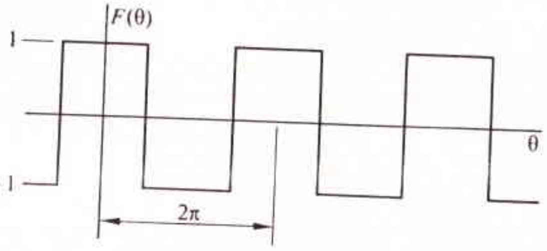


Fig. 12.2–1 Plot of $F(\theta)$ vs. θ.

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DOCUMENTACION Y BIBLIOTECA
MONTEVIDEO - URUGUAY

There are only *two* basic techniques for demodulating [or extracting $f(t)$ from] FM signal. The first technique places an FM modulator in the return branch a feedback amplifier. The resultant demodulator, referred to as the phase-locked op (PLL), makes use of the fact that a feedback amplifier with sufficient loop gain erforms in its forward branch the inverse operation of that which is performed in its turn branch. Since the PLL is somewhat specialized in its operation, we shall defer s explanation to Section 12.7.

The second and more general technique for accomplishing FM demodulation is hat of first placing the amplitude-limited FM signal through a differentiating net-ork which produces an envelope modulation proportional to the instantaneous equency of the FM signal, and then placing the resultant amplitude-varying. equency-modulated signal through an amplitude demodulator which extracts a gnal proportional to $\omega_i(t)$. Since $\omega_i(t) = \omega_0 + \Delta\omega f(t)$, once $\omega_i(t)$ is extracted. (t) may be obtained either by subtracting the ω_0-term or by removing it by placing (t) through a high-pass (capacitive) coupling network. A block diagram of this asic frequency demodulator (or discriminator, as it is also called) is shown in Fig. 2.2–2. More than 99.9% of all the frequency demodulators in existence today mploy some variation of the demodulation technique illustrated by the block iagram.

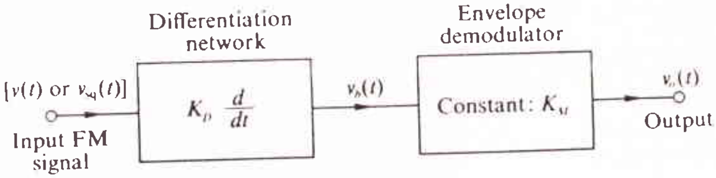


Fig. 12.2–2 Block diagram of the basic frequency demodulator.

To see more quantitatively how the demodulator of Fig. 12.2–2 performs, let us determine an expression for its output when each of the signals given by Eq. (12.2–1) is applied at its input. With $v(t)$ applied, the output of the differentiation network is given by

$$v_b(t) = \overbrace{-K_DA[\omega_0 + \Delta\omega f(t)]}^{\text{envelope proportional to } \omega_i(t)} \sin\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right], \tag{12.2–2}$$

with the result that the output of the envelope demodulator has the desired form

$$v_o(t) = AK_DK_M[\omega_0 + \Delta\omega f(t)] = AK_DK_M\omega_i(t). \tag{12.2–3}$$

where K_D is the constant of the differentiation network and K_M is the constant of the amplitude demodulator. On the other hand, with $v_{sq}(t)$ applied, the output of the differentiation network is given by

$$v_b(t) = K_DB[\omega_0 + \Delta\omega f(t)]F'\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right], \tag{12.2–4}$$

---

† This form is obtained by applying the chain rule when differentiating $v_{sq}(t)$.

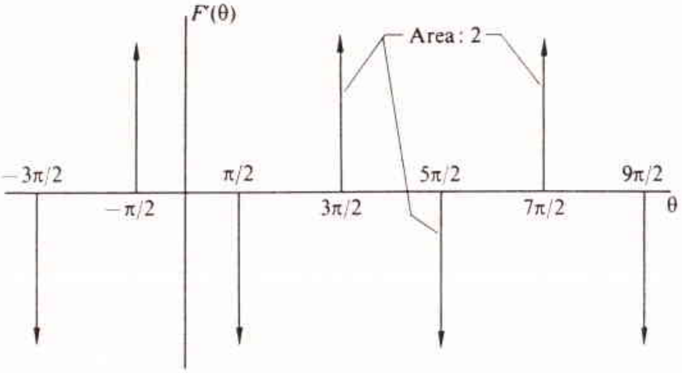


Fig. 12.2–3 Plot of $F'(\theta)$ vs. θ.

where $F'(\theta) = dF(\theta)/d\theta$. A plot of $F'(\theta)$ vs. θ is shown in Fig. 12.2–3. Here we observe that $BK_D[\omega_0 + \Delta\omega f(t)]$ envelope-modulates a high-frequency train of positive and negative impulses of area 2; hence, if the envelope demodulator extracts the envelope, $v_o(t)$ is again given by

$$v_o(t) = K_DK_MB[\omega_0 + \Delta\omega f(t)]. \tag{12.2–5}$$

The value of K_M in this case is in general, but not always, different from K_M given in Eq. (11.2–3), since most amplitude demodulators are waveform dependent.

It is now apparent that to implement the frequency demodulator of Fig. 12.2–2 we require a differentiation network and an amplitude demodulator; hence in the remainder of this section and in the next few sections the physical implementation of these blocks will be considered. Specifically, in the remainder of this section we shall consider the theoretical restrictions imposed on the input FM signal when each of the three amplitude demodulators discussed in Chapter 10 is employed in the frequency demodulator of Fig. 12.2–2. In particular, we shall obtain a curve of the maximum permissible deviation ratio $D = \Delta\omega/\omega_0$ as a function of modulation index $\beta = \Delta\omega/\omega_m$ for which theoretically undistorted demodulation of the input FM signal is possible.

We shall then, in subsequent sections, consider the three different ways in which the differentiation network can be realized or closely approximated. For each of the three cases we shall first obtain the theoretical limits on the operation of the differentiator and then proceed to implement it with a physical network. We shall then combine the differentiator with one of the amplitude demodulators discussed in Chapter 10 to obtain a complete frequency demodulator.

Note that, since there are three basic methods for envelope demodulation and three basic ways to make a differentiating circuit, there are nine different frequency demodulators of the form shown in Fig. 12.2–2. We shall explore only the most practical of these and leave the study of the remainder to the interested reader.

chronous Detection

accomplish the synchronous detection of a differentiated FM signal we multiply given by Eq. (12.2–2) or Eq. (12.2–4) by the undifferentiated FM reference signal $_1 \sin[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta]$, as shown in Fig. 12.2–4, and pass the resultant signal ough a low-pass filter. Where $v_b(t)$ is given by Eq. (12.2–2) and the input FM signal nusoidal, the multiplier output of the synchronous detector has the form

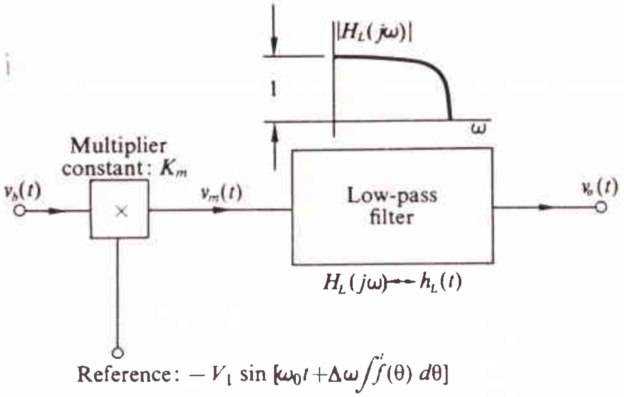

$|H_L(j\omega)|$
1
ω
Multiplier constant: K_m
$v_b(t)$
×
$v_m(t)$
Low-pass filter
$v_o(t)$
$H_L(j\omega) \leftrightarrow h_L(t)$
Reference: $-V_1 \sin[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta]$

Fig. 12.2–4 Synchronous detector for an FM signal.

$$v_m(t) = \overbrace{+ AK_DK_M[\omega_0 + \Delta\omega f(t)]}^{v_{m1}(t)}$$

$$\overbrace{- AK_DK_M[\omega_0 + \Delta\omega f(t)] \cos\left[2\omega_0 t + 2\Delta\omega \int^t f(\theta)\, d\theta\right]}^{v_{m2}(t)}, \qquad (12.2\text{–}6)$$

ere $K_M = KV_1/2$. The multiplier output contains a low-frequency component $t)$ and a high-frequency component $v_{m2}(t)$, as is shown in Fig. 12.2–5. If the spectra hese two terms do not overlap, then $v_{m1}(t)$ may be extracted by the low-pass filter wn in Fig. 12.2–4 to yield

$$v_o(t) = AK_DK_M[\omega_0 + \Delta\omega f(t)] * h_L(t), \qquad (12.2\text{–}7)$$

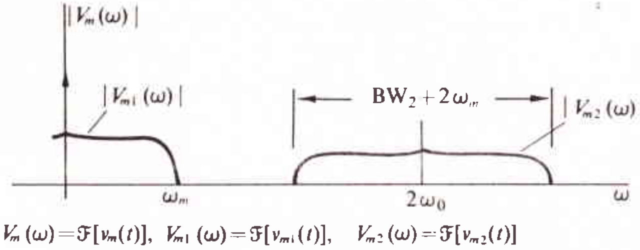


Fig. 12.2–5 Spectrum of $v_m(t)$ plotted vs. ω.

where $h_L(t)$ is the impulse response of the low-pass filter. If the filter is sufficiently flat at $H_L(0) = 1$ over the band of frequencies $0 \le \omega \le \omega_m$, then Eq. (12.2–7) simplifies to the desired form

$$v_o(t) = AK_DK_M[\omega_0 + \Delta\omega f(t)]. \qquad (12.2\text{–}8)$$

The ability to separate $v_{m1}(t)$ from $v_{m2}(t)$ by filtering places a theoretical limit on the maximum values of $\Delta\omega$ and ω_m relative to ω_0 which can be demodulated by following the differentiation network by a synchronous detector. As we shall now show, these limits can be expressed as a plot of the maximum permissible deviation ratio $D = \Delta\omega/\omega_0$ versus $\beta = \Delta\omega/\omega_m$. We shall obtain a similar plot for the average envelope detector and the peak envelope detector which will indicate that synchronous detection permits the largest possible deviation ratio for any given value of β.

To determine under what conditions the spectral components of $V_m(\omega)$ do not overlap, we note that the upper spectral component $V_{m2}(\omega)$ is the Fourier transform of the product of $\omega_0 + \Delta\omega f(t)$ (which is band-limited to ω_m) and an FM signal centered at $2\omega_0$ with a frequency deviation $\Delta\omega_2 = 2\,\Delta\omega$, a modulation index $\beta_2 = 2\beta$, and a bandwidth BW_2. Since multiplication in the time domain corresponds to convolution in the frequency domain, the spectrum of $V_{m2}(\omega)$ has a bandwidth of $\mathrm{BW}_2 + 2\omega_m$.† Therefore, separation of $v_{m1}(t)$ from $v_{m2}(t)$ by filtering is possible provided that

$$\frac{\mathrm{BW}_2}{2} + 2\omega_m \le 2\omega_0. \qquad (12.2\text{–}9)$$

or equivalently,

$$D = \frac{\Delta\omega}{\omega_0} \le \frac{1}{\dfrac{1}{\beta} + \dfrac{\mathrm{BW}_2}{2\,\Delta\omega_2}} = D_{\max}. \qquad (12.2\text{–}10)$$

For any value of β a conservative estimate of $\mathrm{BW}_2/2\,\Delta\omega_2$ can be obtained from Fig. 11.1–4; thus $D_{\max}$, the maximum deviation ratio for which synchronous detection (or any form of detection, for that matter) is possible, may be plotted as a function of β. For example, if $\beta = 5$, then $\beta_2 = 2\beta = 10$ and $\mathrm{BW}_2/\Delta\omega_2 = 1.4$. Therefore, $D_{\max} = 0.625$.

A plot of $D_{\max}$ vs. β for the synchronous detector is shown in Fig. 12.2–6. As $\beta \to \infty$, $D_{\max} \to 1$; hence the curve has a known limiting value. If the input FM signal has a modulation index β and a deviation ratio D which determine a point lying under the curve of Fig. 12.2–6, then it is possible to demodulate it with a frequency demodulator employing a synchronous detector. However, if the point lies close to the curve, the complexity of the filter at the output of the synchronous detector will increase significantly.

---

† Convolution of two band-limited spectra produces a spectrum which is also band-limited, with a bandwidth equal to the sum of the two original bandwidths.

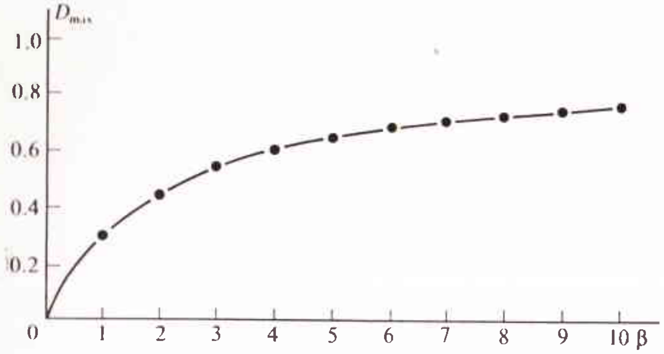


Fig. 12.2–6 Plot of $D_{\max}$ vs. β for which synchronous demodulation of an FM signal is possible.

For the case where $v_b(t)$ given by Eq. (12.2–4) is applied to the synchronous detector, the effect of multiplying $F'[\omega_0\tau(t)]$ by $-\sin\omega_0\tau(t)$ is the conversion of the train of alternating positive and negative impulses into a train of all positive impulses. However, since this full-wave rectification can be performed in a more straightforward fashion with diode networks, synchronous detection is not employed in conjunction with hard-limited FM signals.

To develop the reference signal for the synchronous detector, the input FM signal may be placed through a network which shifts its phase by $\pi/2$. Another method, the one most often used in practice, is to shift the output of the differentiation network by $-\pi/2$ and use the undifferentiated input FM signal directly as the reference.

Average Envelope Detection

When $v_b(t)$ given by Eq. (12.2–2) is placed through an average envelope detector of the orm shown in Fig. 10.1–6, $v_a(t)$ may be written as the product of $v_b(t)$ and a switching unction $S_{\mathrm{FM}}(t)$ which is a unit amplitude square wave similar to the one shown in ig. 10.1–7 except that it is periodic in $\tau(t)$ rather than t. Thus $S_{\mathrm{FM}}(t)$ may be expanded n a Fourier series, with the result that $v_a(t)$ may be written as follows (cf. Eq. 10.1–11)

$$v_a(t) = \overbrace{\frac{K_DA}{\pi}[\omega_0 + \Delta\omega f(t)]}^{v_{a1}(t)} + \overbrace{\frac{K_DA}{2}[\omega_0 + \Delta\omega f(t)]\sin\omega_0\tau(t)}^{v_{a2}(t)}$$
$$+ \text{ higher harmonic AM–FM signals.} \tag{12.2–11}$$

If $v_{a1}(t)$ can be extracted from $v_{a2}(t)$ by the output filter of Fig. 10.1–6, then

$$v_o(t) = \frac{K_DA}{\pi}[\omega_0 + \Delta\omega f(t)] * h_L(t). \tag{12.2–12}$$

n addition, if the low-pass filter passes $f(t)$ in an undistorted fashion, $v_o(t)$ simplifi o the desired form

$$v_o(t) = K_MK_DA[\omega_0 + \Delta\omega f(t)], \tag{12.2–13}$$

here $K_M = H_L(0)/\pi$.

In this case, in order to separate $v_{a1}(t)$ from $v_{a2}(t)$ by filtering, the spectrum of the amplitude-modulated FM wave centered at ω_0 with a bandwidth $\mathrm{BW}_1 + 2\omega_m$ (where BW_1 is the bandwidth of the input FM signal) must not overlap with the low-frequency spectrum of $f(t)$, which extends to ω_m. To ensure this separation, we require that

$$2\omega_m + \frac{\mathrm{BW}_1}{2} \le \omega_0, \tag{12.2–14}$$

or equivalently,

$$D \le \frac{1}{\dfrac{2}{\beta} + \dfrac{\mathrm{BW}_1}{2\,\Delta\omega}}. \tag{12.2–15}$$

With the aid of Fig. 11.1–4, $D_{\max}$ may be plotted as a function of β, as shown in Fig. 12.2–7. By comparing Fig. 12.2–7 with Fig. 12.2–6 (which is replotted with dashed lines in Fig. 12.2–7), we observe that, since the synchronous detector "ripple" term was centered at $2\omega_0$ while the average detector "ripple" term is centered at only ω_0, the synchronous detector is less restrictive on the parameters of the input FM signal than is the average envelope detector. However, if the parameters of the input FM signal determine a point which lies under the curve of Fig. 12.2–7, average envelope detection may be employed instead of the more complex synchronous detection.

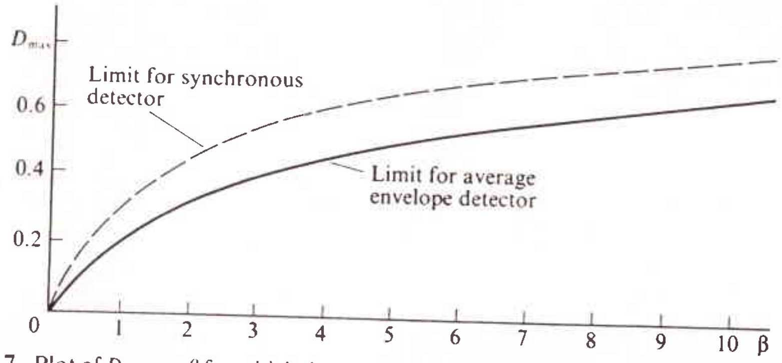


Fig. 12.2–7 Plot of $D_{\max}$ vs. β for which demodulation of an FM signal is possible with an average envelope detector.

When $v_b(t)$ given by Eq. (12.2–4) is placed through the average envelope detector of Fig. 10.1–6, $v_a(t)$ has the form

$$v_a(t) = K_DB[\omega_0 + \Delta\omega f(t)]F_p'[\omega_0\tau(t)], \tag{12.2–16}$$

where $F_p'(\theta)$ is shown in Fig. 12.2–8. Equation (12.2–16) is a direct result of the fact that the half-wave rectifier in the average envelope detector removes the negative impulses from $F'(\theta)$.

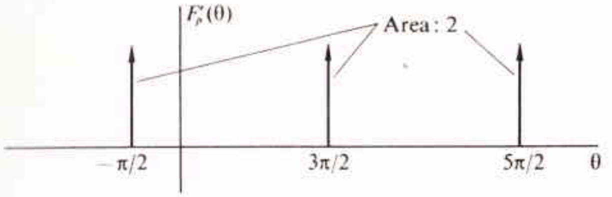


Fig. 12.2-8 Plot of $F_p'(\theta)$ vs. θ.

It is apparent that $F_p'[\omega_0\tau(t)]$ may be expanded in a Fourier series in $\omega_0\tau(t)$, here the average value of the series is given by $1/\pi$; thus the low-frequency com- onent of $v_a(t)$ has the form $(K_DB/\pi)[\omega_0 + \Delta\omega f(t)]$. If this component of $v_a(t)$ can e separated from the fundamental AM–FM signal component of $v_a(t)$ by the output w-pass filter, then

$$v_o(t) = \frac{K_DB}{\pi}[\omega_0 + \Delta\omega f(t)] * h_L(t) \tag{12.2–17}$$

r

$$v_o(t) = K_DK_MB[\omega_0 + \Delta\omega f(t)], \tag{12.2–18}$$

here $K_M = H_L(0)/\pi$ for the case where $h_L(t)$ passes $f(t)$ undistorted. It is interesting note that K_M for the average envelope detector is the same for both impulselike d sinusoidal inputs.

Again in this case, to ensure the validity of Eq. (12.2–17) or Eq. (12.2–18), the ectrum of $f(t)$ must not overlap with the spectrum of the component of $v_a(t)$ centered ω_0. This requirement is, of course, identical with that which ensures the separation f $v_{a1}(t)$ from $v_a(t)$ given in Eq. (12.2–11); hence for an unfiltered FM signal the curve f Fig. 12.2–7 again determines the maximum value of the deviation ratio for any ven β for which average envelope detection may be employed in the frequency emodulator.

If a full-wave rectifier replaces the half-wave rectifier of Fig. 10.1–7, the output mplitude is, of course, increased by a factor of 2. More important, however, as the terested reader should verify, the $D_{\max}$-vs.-β curve becomes identical to the one btained with a synchronous detector. The basic reason for this is that with full-wave ectification the component of $v_a(t)$ centered about ω_0 is reduced to zero.

eak Envelope Detection

peak envelope detector either greatly distorts a train of narrow pulses or functions sentially as an average detector when such pulses are applied at its input; hence is seldom employed in the frequency demodulation of unfiltered square-wave FM gnals.

On the other hand, when the limited FM signal is band-limited and has the form given by Eq. (12.2–1a), the peak envelope detector may be employed. However, when it is employed, as we recall from Sections 10.3 and 10.5, the carrier frequency must be much greater than ω_m, usually by a factor of 100 or more, for its correct operation (i.e., to keep the ripple small and to prevent failure-to-follow distortion). For the case of FM signals it is the instantaneous frequency $\omega_i(t)$ rather than the carrier frequency which gives rise to ripple at the detector output; hence, when the peak envelope detector is employed in the frequency demodulator, we must require

$$|\omega_i(t)|_{\min} = \omega_0 - \Delta\omega > 100\omega_m, \tag{12.2–19}$$

or equivalently,

$$D < \frac{\beta}{100 + \beta}. \tag{12.2–20}$$

A plot of $D_{\max} = \beta/(100 + \beta)$ is shown in Fig. 12.2–9 along with the plot of Fig. 12.2–6 (dashed lines) on the same set of coordinates. Again we observe that the synchronous detector is far less restrictive on the parameters of the input FM signal than is the peak envelope detector. Since $D_{\max} \to 1$ as $\beta \to \infty$ for all three amplitude demodulators, it makes little or no difference theoretically which envelope detector is used where the modulation index is high. Simplicity of construction in practical circuits usually dictates the use of the peak or average envelope detector in place of the synchronous detector.

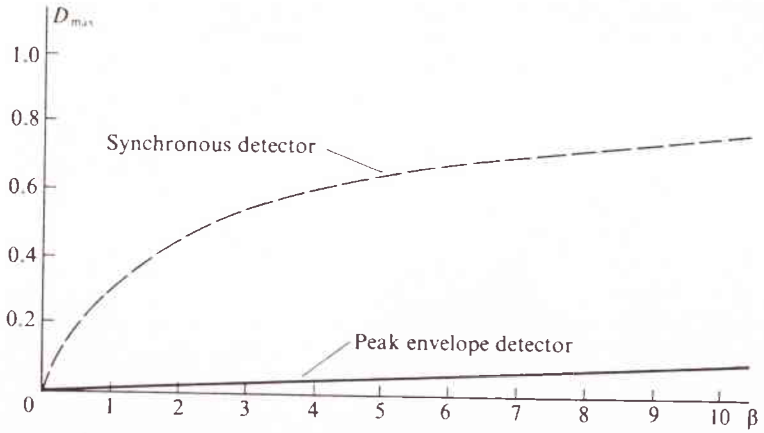


Fig. 12.2–9 Plot of $D_{\max}$ vs. β for the peak envelope detector.

Note that, if the FM signal has been distorted somewhere in the system prior to the demodulator, for example by a limiter, and then filtered to reclaim $v(t)$, $D_{\max}$ must be lower than the top curve of Fig. 11.2–3. However, since this filtering curve is more restrictive than the amplitude demodulator limiting curves on $D_{\max}$ given in Figs. 12.2–6 and 12.2–7, D is normally limited by the bandpass filters in the system and not by the amplitude demodulator in the discriminator.

To put things into proper perspective, one should recall that ordinary entertainment-type FM broadcasting has $\Delta f = 75$ kHz and $f_m = 15$ kHz and uses an IF of 10.7 MHz. Hence $D_{\max} = 0.0070$ and $\beta = 5$. Thus a well-designed peak detector will never pose a problem for such a system, since $D_{\max}$ can never occur with β less than 5; however, Eq. (12.2–20) is satisfied down to $\beta = 0.7$.

Only in special wideband systems will the restrictions of this section become important.

12.3 DIRECT DIFFERENTIATION— THE CLARKE-HESS FREQUENCY DEMODULATOR

Three methods exist by which the differentiation of a signal may be accomplished:

1. Direct differentiation.
2. Frequency-domain differentiation.
3. Time-delay differentiation.

In this section and in subsequent sections each of these differentiation methods is investigated theoretically, implemented practically, and combined with a practical envelope detector to yield a frequency demodulator (or discriminator) of the basic form shown in Fig. 12.2–2. In this section we consider the direct differentiator.

A direct differentiator is a device which performs the operation of differentiation exactly and thus places no theoretical limitations on the signal at its input. When such a device is employed in a frequency demodulator, all the limitations on the incoming FM signal arise from the amplitude demodulator. An example of a direct differentiator would be a circuit, employing a single capacitor C, for which the input is the capacitor voltage and the output is the capacitor current.

A frequency demodulator employing a capacitor as a direct differentiator and employing an average envelope detector of the form shown in Fig. 10.2–5 is illustrated in Fig. 12.3–1. A model for the circuit of Fig. 12.3–1 in which both D and the emitter-base junction of Q are modeled by ideal diodes in series with batteries of value V_0 is shown in Fig. 12.3–2. For proper circuit operation, V_0 must be small in comparison with the maximum value of $v_i(t)$ (as we saw in Section 10.2); i.e., germanium diodes and transistors should be employed. Hence in the following analyses we assume $V_0 \approx 0$.

With $V_0 \approx 0$, the two ideal diodes in parallel are indistinguishable from a short circuit; thus the capacitor voltage is given by

$$v_C(t) = v_i(t) - V_{CC}, \tag{12.3–1}$$

with the result that the capacitor current has the form

$$i_i(t) = C\frac{dv_i(t)}{dt}. \tag{12.3–2}$$

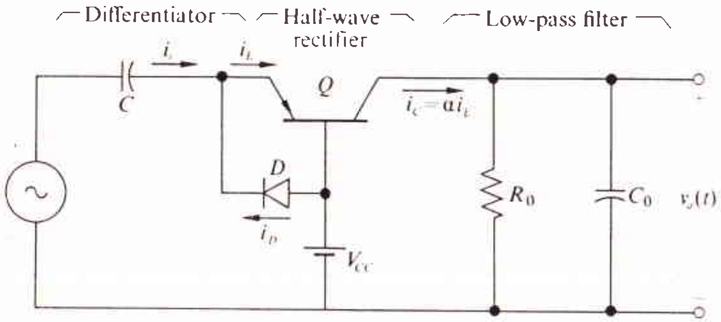


Fig. 12.3–1 Frequency demodulator employing a direct differentiator and an average envelope detector.

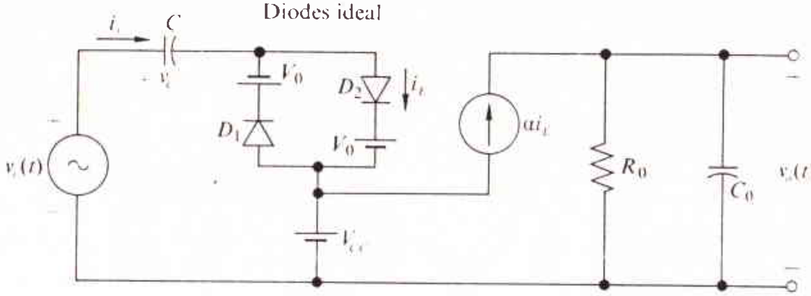


Fig. 12.3–2 Model for the discriminator of Fig. 12.3–1.

The capacitor current is a direct derivative of $v_i(t)$ with the differentiation constant $K_D = C$. (In practice, a finite generator impedance term and a finite diode and transistor loss term combine as r_Σ to make the differentiation approximate instead of exact. If $\omega_0 C r_\Sigma < 0.1$, then the operation is within 1% of pure differentiation. If $r_\Sigma = 50\,\Omega$ and $\omega_0 = 10^7$ rps, then C must be less than 200 pF.)

If $v_i(t) = v(t)$, where $v(t)$ is given by Eq. (12.2–1a), then

$$i_i(t) = -AC[\omega_i(t)] \sin\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta\right]. \tag{12.3–3}$$

Only the positive half of $i_i(t)$ flows into D_2; hence $i_E(t)$ is a half-wave rectified version of $i_i(t)$. The average value of this rectified waveform is extracted by placing $\alpha i_E(t)$ through the R_0C_0 low-pass filter. Consequently, if $D = \Delta\omega/\omega_0$ and $\beta = \Delta\omega/\omega_m$ for $v(t)$ determine a point below the curve of Fig. 12.2–7 and the low-pass R_0C_0 filter extracts the low-frequency component of $\alpha i_E(t)$, then $v_o(t)$ may be written as follows (cf. Eq. 12.2–12):

$$v_o(t) = \frac{\alpha AC\omega_i(t) * h_L(t)}{\pi}. \tag{12.3–4}$$

where $h_l(t)$ is the impulse response of the low-pass filter. For the case where the filter passes $\omega_i(t) = \omega_0 + \Delta\omega f(t)$ in an undistorted fashion (the filter is flat for $0 \leq \omega \leq \omega_m$), Eq. (12.3–4) simplifies to the desired form

$$v_o(t) = \frac{\alpha ACR_0}{\pi}\omega_i(t)$$

$$= \frac{\alpha ACR_0}{\pi}[\omega_0 + \Delta\omega f(t)]. \qquad (12.3\text{–}5)$$

A plot of v_o versus ω_i is shown in Fig. 12.3–3 for this case. From the figure we observe that the output voltage is linearly related to $\omega_i(t)$ up to the frequency for which $v_o = V_{CC} + V_0 \approx V_{CC}$ and Q saturates. For any given $\omega_i(t)$, the slope of the characteristic shown in Fig. 12.3–3 must be chosen sufficiently small so that Q does not saturate. For example, if $A = 10$ V, if $\omega_i(t) = 10^8[1 + \frac{1}{2}f(t)]$ rad/sec, if $C = 50$ pF, and if $V_{CC} = 15$ V, then R_0 must be chosen less than

$$\frac{\pi V_{CC}}{\alpha AC\omega_{i_{\max}}} = \frac{(3.14)(15)}{(50 \times 10^{-12})(1.5 \times 10^8)(10)} = 628\ \Omega.$$

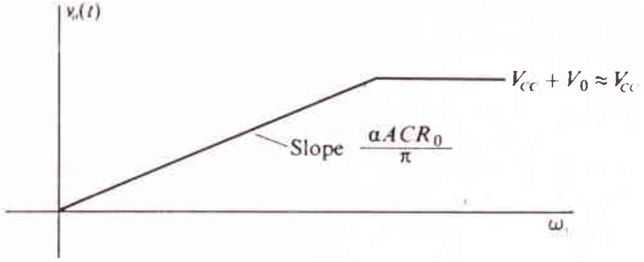


Fig. 12.3–3 Plot of the discriminator characteristic for the frequency demodulator of Fig. 12.3–1

Clarke-Hess Frequency Demodulator†

The circuit shown in Fig. 12.3–2 functions best when $\Delta\omega$ is a reasonable fraction of ω_0. However, when $\Delta\omega \ll \omega_0$, significant advantages over the circuit of Fig. 12.3–2 result when a balancing branch, which reduces v_o to zero when $\omega_i = \omega_0$, is added to the basic frequency demodulator. A circuit containing such a branch is shown in Fig. 12.3–4. The balancing branch in the circuit subtracts a term proportional to the input envelope from $v_o(t)$. In particular, if $V_0 \approx 0$, a term of value $\alpha AR_0/\pi R$ (cf. Eq. 10.2–11) is subtracted from the output, with the result that $v_o(t)$ takes the form

$$v_o(t) = \frac{\alpha ACR_0}{\pi}\left[\omega_0 + \Delta\omega f(t) - \frac{1}{RC}\right]. \qquad (12.3\text{–}6)$$

---

† K. K. Clarke and D. T. Hess, "Demodulator for Frequency Modulated Signals," U.S. Patent 3292093, December 13, 1966.

However, if $1/RC$ is adjusted to equal ω_0, then $v_o(t)$ simplifies to

$$v_o(t) = \left(\frac{\alpha ACR_0}{\pi}\right)\Delta\omega f(t). \qquad (12.3\text{–}7)$$

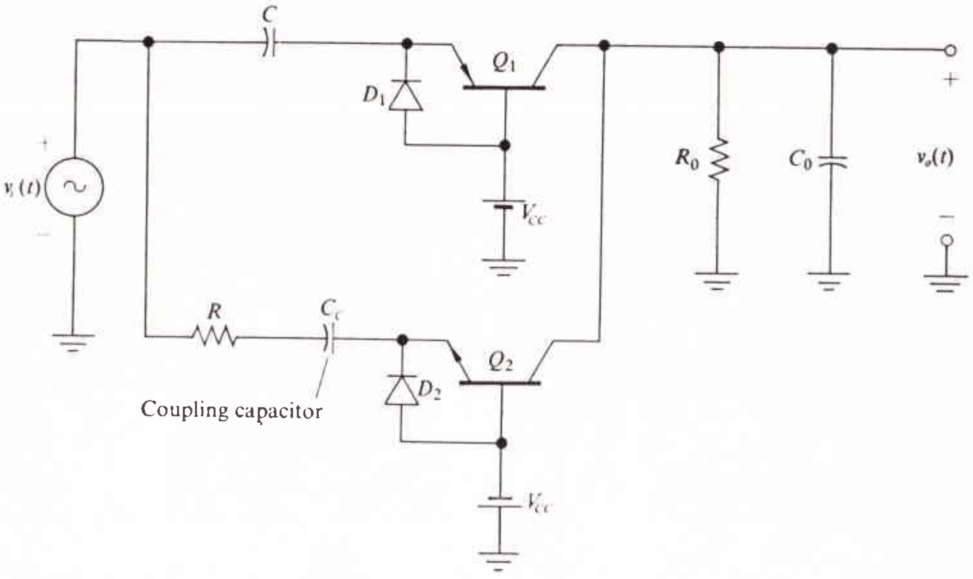


Fig. 12.3–4 Clarke–Hess frequency demodulator.

The transfer function of $v_o(t)$ given by Eq. (12.3–7) vs. $\Delta\omega f(t)$ is shown in Fig. 12.3–5. Here we observe that v_o is linearly related to $\Delta\omega f(t)$ provided that neither Q_1 nor Q_2 saturates.

The advantages of balancing out the carrier component from the output by choosing $RC = 1/\omega_0$ are numerous. First, the balancing suppresses small amplitude

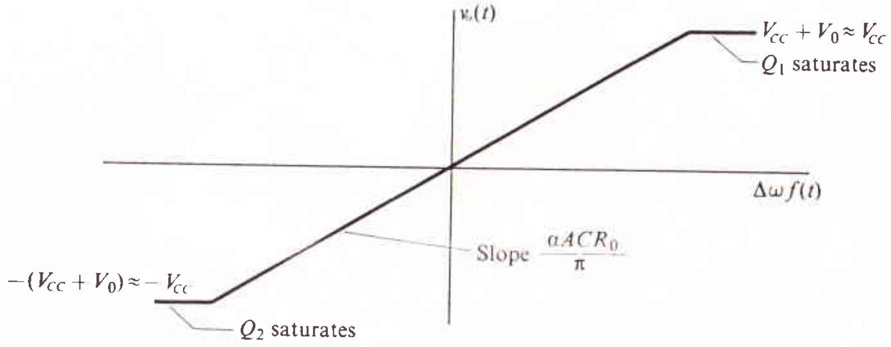


Fig. 12.3–5 Discriminator characteristic for the Clarke–Hess frequency demodulator.

variations which might sneak by the limiter. For example, if

$$A = A_0[1 + \varepsilon(t)],$$

where $|\varepsilon(t)| \ll 1$ is a small disturbance on the envelope A, then at the output of the circuit of Fig. 12.3–4

$$v_o(t) = \frac{\alpha A_0 C R_0}{\pi}\{\Delta\omega f(t)[1 + \varepsilon(t)]\}. \tag{12.3–8}$$

Here we see that a corresponding small disturbance appears at the output which decreases in direct proportion to $f(t)$. On the other hand, with $A = A_0[1 + \varepsilon(t)]$, the output of the single-ended circuit of Fig. 12.3–1 is given by

$$\begin{aligned} v_o(t) &= \frac{\alpha A_0 C R_0}{\pi}[\omega_0 + \Delta\omega f(t) + \omega_0\varepsilon(t) + \Delta\omega f(t)\varepsilon(t)] \\ &\approx \frac{\alpha A_0 C R_0 \omega_0}{\pi}\left[1 + \frac{\Delta\omega}{\omega_0} f(t) + \varepsilon(t)\right]. \end{aligned} \tag{12.3–9}$$

In this case the output disturbance is independent of $f(t)$ and, if $\Delta\omega/\omega_0$ is small, the disturbance might easily dominate $f(t)$. For this reason alone, almost all high-quality FM demodulators are balanced in some fashion or another. [Note that subtracting a constant which is independent of A from the output does not produce the same effect as balancing, even though $v_o(t) = 0$ for $\omega_i(t) = \omega_0$.]

The balancing branch also provides a means of developing an error signal for the automatic frequency control (AFC) in an FM receiver or the frequency stabilizing circuit of an FM generator. If $1/RC$ is adjusted to equal ω_0 (which is usually the center frequency of the IF strip in an FM receiver) and the FM signal passing through the IF strip drifts in frequency by $\delta\omega$, then the output of the circuit of Fig. 12.3–4 is given by

$$v_o(t) = \frac{\alpha A C R_0}{\pi}[\Delta\omega f(t) + \delta\omega]. \tag{12.3–10}$$

If $v_o(t)$ is placed through a sufficiently low-pass filter to remove the $f(t)$ component, i.e., the demodulated signal, then the filter output contains a dc signal directly proportional to the frequency drift. This signal can then be amplified and applied as a frequency control to the local oscillator (which in this case is an FM generator) to return the mixer output signal to the center of the IF strip.

Third, balancing permits us to increase greatly the sensitivity of $v_o(t)$ to $f(t)$. As can be seen from Fig. 12.3–5, the slope of the $\Delta\omega f(t)$-$v_o(t)$ characteristic can be increased to the point where variations in $f(t)$ between $+1$ and -1 cause corresponding variations in $v_o(t)$ between $+V_{CC}$ and $-V_{CC}$; that is, a slope of $V_{CC}/\Delta\omega$ is possible. On the other hand, without balancing, the characteristic of Fig. 12.3–3 prevails. This characteristic can have a maximum slope (without saturating the transistor) of

$$\frac{V_{CC}}{\omega_{i\max}} = \frac{V_{CC}}{\omega_0 + \Delta\omega}.$$

Thus with balancing the sensitivity may be increased by a factor of

$$\frac{\omega_0 + \Delta\omega}{\Delta\omega} = 1 + \frac{\omega_0}{\Delta\omega},$$

which is quite significant for most practical circuits in which $\omega_0 \gg \Delta\omega$.

Example 12.3–1 Choose values for the various components in the circuit of Fig. 12.3–4 such that the circuit produces a 2 V peak-to-peak output when driven by an FM signal of the form

$$v_i(t) = (10\text{ V})\cos\left[10^7 t + 10^5 \int^t \cos 2 \times 10^4 t\right].$$

The available power supplies are $\pm V_{CC} = \pm 12$ V.

Solution. We first note that $\omega_0 = 10^7$ rad/sec, $\Delta\omega = 10^5$ rad/sec, and $\omega_m = 2 \times 10^4$ rad/sec; thus $\beta = 5$ and $D = 1/100$. Therefore, since D and β determine a point well below the curve of Fig. 12.2–6, we observe that demodulation is possible without requiring an elaborate output low-pass filter.

We next select two germanium transistors and two germanium diodes with frequency limits well above 10 MHz. In addition, we select $R = 5\text{ k}\Omega$ as a compromise between a large output signal and an increased value of V_0. As we saw in Section 10.2, a small value for R increases the effective size of V_0. On the other hand, a large value of R decreases the sensitivity of the demodulator.

With $R = 5\text{ k}\Omega$ and $\omega_0 = 10^7$ rad/sec we obtain $C = 1/\omega_0 R = 20$ pF. To obtain a 2 V peak-to-peak output signal, the slope of the characteristic of Fig. 12.3–5 must be adjusted to equal (1 V)/10^5 rad/sec; thus R_0 must have the value

$$R_0 = \frac{1\text{ V}}{10^5\text{ rad/sec}}\,\frac{\pi}{\alpha A C} = 157\text{ k}\Omega.$$

In addition, to pass the modulation information, $R_0 C_0$ must be less than or equal to $1/\omega_m$. If we choose $R_0 C_0 = 1/\omega_m$, then $C_0 = 318$ pF. (Actually, for this value of $R_0 C_0$ the output baseband signal suffers a 3 dB attenuation. To avoid this attenuation C_0 must be reduced; this leads to an increase in output ripple.)

Finally, a choice of $C_C = 2000$ pF ensures that the coupling capacitor is an effective short circuit in comparison with R at the carrier frequency ($X_C = R/100$).

If an FM square wave is applied to the frequency demodulator shown in Fig. 12.3–4, the circuit functions in exactly the same fashion as it did for a sinusoidal FM signal, provided that RC is readjusted so that

$$RC = \pi/2\omega_0 . \tag{12.3–11}$$

As was pointed out in Section 12.2, when an FM square wave is placed through a differentiator and an average envelope detector, the output is identical with that produced by a sinusoidal FM signal with the same amplitude. Hence the contribution to $v_o(t)$ from the upper branch of the circuit of Fig. 12.3–4 is unchanged when an FM square wave is applied. On the other hand, when an FM square wave of amplitude B is applied to the balancing branch of the demodulator, a component of value $BR_0/2R$ is subtracted from v_o, as the reader should verify. Thus, with an FM square wave applied to the frequency demodulator of Fig. 12.3–4, $v_o(t)$ is given by

$$v_o(t) = \frac{\alpha BCR_0}{\pi}\left[\omega_0 + \Delta\omega f(t) - \frac{\pi}{2RC}\right], \tag{12.3–12}$$

which reduces to

$$v_o(t) = \frac{\alpha BCR_0}{\pi}\Delta\omega f(t) \tag{12.3–13}$$

provided that $RC = \pi/2\omega_0$. Clearly, $v_o(t)$ given by Eq. (12.3–13) is identical with $v_o(t)$ given by Eq. (12.3–7) for the case where the FM square wave and the sinusoidal FM signal have the same amplitude.

With a square-wave input to the frequency demodulator of Fig. 12.3–4, a much larger ripple component appears at the output than with a corresponding sinusoidal input. This is a direct result of the narrow current pulses of area $2BC$ which flow through the upper branch of the frequency demodulator into the output capacitor C_0. In particular, the ripple on $v_o(t)$ from this component alone has a value of $2BC/C_0$. Of course, this ripple can be removed by subsequent stages of filtering; however, it still adds directly to $v_o(t)$. This direct addition limits the maximum amplitude of the signal component of $v_o(t)$, since the signal plus the ripple must not saturate the transistors of the demodulator. In particular, the maximum signal component of $v_o(t)$ must be kept less than $V_{CC} + V_0 - 2BC/C_0$.

In the practical circuit the finite value of r_Σ caused by the generator and transistor-diode impedances will cause the "differentiator" branch current pulses to take the shape of decaying exponential pulses of initial amplitude $2B/r_\Sigma$ rather than impulses of infinite amplitude. In either case the pulse area will be $2BC$. In the practical case one must make sure that $5r_\Sigma C$ is smaller than a half-period of the highest instantaneous frequency $f_0 + \Delta f$ in order that the pulse caused by one step is completed before the next step occurs. This requires that

$$C \leq \frac{1}{10(f_0 + \Delta f)r_\Sigma}. \tag{12.3–14}$$

12.4 FREQUENCY-DOMAIN DIFFERENTIATION—THE SLOPE DEMODULATOR

A frequency-domain differentiator is a linear network whose transfer function $H(j\omega)$ has a linearly sloping magnitude over the band of frequencies occupied by the input FM signal. Such a network, as was shown in Section 11.3, functions as a differentiator and converts variations in the instantaneous frequency at its input to envelope variations at its output. For example, if the network has the transfer function shown in Fig. 11.3–2 and if $v(t)$ given by Eq. (12.2–1a) is applied at its input, the network output is given by (cf. Eq. 11.3–20)

$$v_b(t) = A\frac{\omega_i(t - t_0) - \omega_1}{\omega_\alpha}\cos\left[\omega_0 t + \Delta\omega\int^{t-t_0} f(\theta)\,d\theta + \theta(\omega_0)\right]. \tag{12.4–1}$$

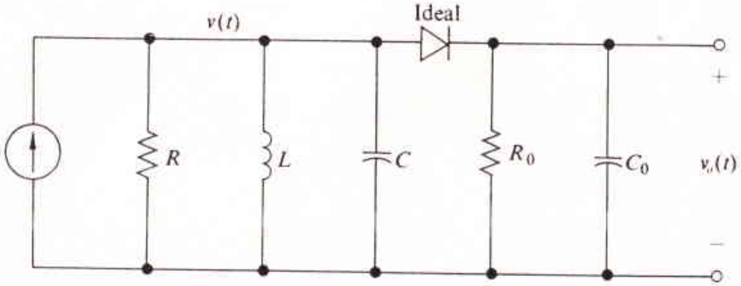
Hence at the output of the amplitude demodulator following the frequency-domain differentiator we obtain

$$v_o(t) = AK_DK_M[\omega_0 - \omega_1 + \Delta\omega f(t - t_0)], \tag{12.4–2}$$

where $K_D = 1/\omega_\alpha$ is the slope of the $|H(j\omega)|$-vs.-ω characteristic and K_M is the constant of the envelope demodulator. It is apparent that a frequency-domain differentiator followed by an amplitude demodulator does indeed produce a signal proportional to $f(t)$ [to be more precise, $f(t - t_0)$].

However, because most practical filters have linear characteristics only over a narrow band of frequencies, a frequency demodulator employing a frequency-domain differentiator is best suited for FM signals for which the deviation ratio is small. This frequency restriction also prevents the demodulation of an FM square wave (which has frequency components in the vicinity of ω_0, $3\omega_0$, $5\omega_0$, etc.) unless the square wave is first filtered to extract its fundamental FM component. This filtering, in some cases, can be accomplished by the differentiation network itself.

The simplest practical implementation of the frequency-domain differentiator is a parallel RLC circuit which is tuned so that the input FM signal is centered on the sloping portion of the transfer function. Figure 12.4–1 illustrates such a circuit followed by a peak envelope detector to form a "slope" demodulator.



$$i_i(t) = I_1 \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\,d\theta\right]$$

Fig. 12.4–1 Single-ended slope frequency demodulator.

If we assume initially that $R_0 \gg R$ and $C_0 \ll C$ so that the envelope detector does load the tuned circuit, that $\Delta\omega$ and ω_m have appropriate values, in comparison h the parameters of the tuned circuit, for the "quasi-static" approximation of (11.3–3) to be valid (we shall check these assumptions), that $i_i(t) = I_1 \cos [\omega_0 t +$ $\int^t f(\theta)\, d\theta]$, and that D and β for the input FM signal determine a point below the ve of Fig. 12.2–8, then the envelope variation in $v(t)$, which appears as $v_o(t)$, is given

$$\begin{aligned} v_o(t) &= I_1|Z_{11}[j\omega_i(t)]| \\ &= I_1|Z_{11}[j\omega_0 + j\,\Delta\omega f(t)]|, \end{aligned} \qquad (12.4\text{–}3)$$

ere $Z_{11}(p)$ is the input impedance of the parallel RLC circuit. Figure 12.4–2 icts this transfer characteristic which relates $\omega_i(t)$ and $v_o(t)$.

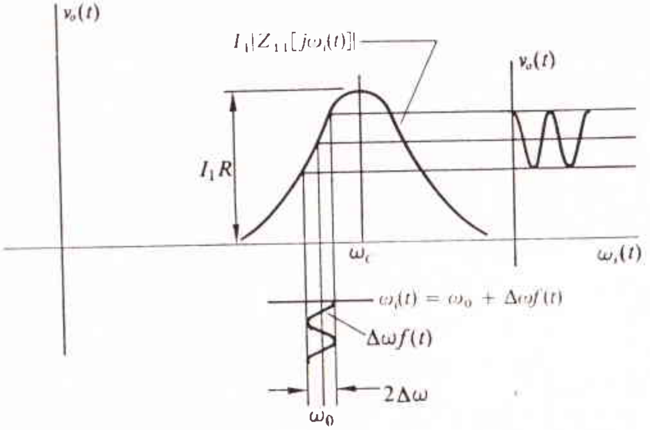


Fig. 12.4–2 Transfer characteristic relating $\omega_i(t)$ and $v_o(t)$.

To design a slope demodulator we must select values for the center frequen , and the bandwidth BW of the tuned circuit of Fig. 12.4–1. In general, we choo ese parameters to achieve the largest possible output signal which is linear lated to $\omega_i(t)$. To see how this choice can best be accomplished, we expand $|Z_{11}($ a Taylor series in ω about ω_0:

$$\begin{aligned} |Z_{11}(j\omega)| &= |Z_{11}(j\omega_0)| + |Z_{11}(j\omega_0)|'(\omega - \omega_0) \\ &\quad + |Z_{11}(j\omega_0)|''\frac{(\omega - \omega_0)^2}{2!} + |Z_{11}(j\omega_0)|'''\frac{(\omega - \omega_0)^3}{3!} + \cdots . \end{aligned} \qquad (12.4$$

is apparent that maximum linearity is achieved if ω_c is chosen such that $|Z_{11}(j$ d as many subsequent derivatives of $|Z_{11}(j\omega)|\big|_{\omega=\omega_0}$ as possible are zero. Th andwidth is then chosen as narrow as possible to maximize $|Z_{11}(j\omega_0)|'$. [A sma andwidth steepens the skirts of $|Z_{11}(j\omega)|$ and thus increases its slope.] However, a nall bandwidth also increases the size of the first nonzero derivative of $|Z_{11}(j\omega)|\,|$,

---

thereby increasing the nonlinear distortion for a fixed $\Delta\omega$. The permissible distortion puts a limit on the minimum value of the bandwidth.

If Q_T for the parallel RLC circuit is greater than 10, then $|Z_{11}(j\omega)|$ can be closely approximated by (cf. Eq. 2.2–9)

$$|Z(j\omega)| = \frac{R}{\sqrt{1 + \left(\dfrac{\omega - \omega_c}{\alpha}\right)^2}}, \qquad (12.4\text{–}5)$$

where $\alpha = 1/2RC = \mathrm{BW}/2$ and $\omega_c = 1/\sqrt{LC}$. For this case

$$\begin{aligned} |Z_{11}(j\omega_0)|' &= \frac{-R\Omega}{\alpha(1 + \Omega^2)^{3/2}}, \\ |Z_{11}(j\omega_0)|'' &= \frac{R(2\Omega^2 - 1)}{\alpha^2(1 + \Omega^2)^{5/2}}, \\ |Z_{11}(j\omega_0)|''' &= \frac{-3R(2\Omega^3 - 3\Omega)}{\alpha^3(1 + \Omega^2)^{7/2}}, \end{aligned} \qquad (12.4\text{–}6)$$

where $\Omega = (\omega_0 - \omega_c)/\alpha$. It is apparent that $|Z_{11}(j\omega_0)|'' = 0$ provided that $\Omega^2 = \frac{1}{2}$ or, equivalently, that

$$\omega_c = \omega_0 \pm \alpha/\sqrt{2}. \qquad (12.4\text{–}7)$$

The plus term of Eq. (12.4–7) corresponds to operation on the lower skirt of $|Z_{11}(j\omega)|$, while the minus term corresponds to operation on the upper skirt. With $\Omega^2 = \frac{1}{2}$, none of the higher derivatives of $|Z_{11}(j\omega)|_{\omega=\omega_0}$ may be set equal to zero. In addition, with Eq. (12.4–7) satisfied,

$$|Z_{11}(j\omega_0)| = \sqrt{\tfrac{2}{3}}R, \qquad |Z_{11}(j\omega_0)|' = \pm\frac{2R}{3\sqrt{3}\alpha},$$

$$|Z_{11}(j\omega_0)|''' = \mp\frac{6R}{\alpha^3\sqrt{3}}\left(\frac{2}{3}\right)^3;$$

hence, $v_o(t)$ may be written (on the assumption that the quasi-static approximation is valid) as

$$\begin{aligned} v_o(t) &= I_1|Z_{11}[j\omega_0 + j\,\Delta\omega f(t)]| \\ &= I_1R\sqrt{\tfrac{2}{3}}\left\{1 + \frac{\sqrt{2}}{3}\frac{\Delta\omega f(t)}{\alpha} \mp 2\sqrt{2}\left(\frac{2}{3}\right)^2\left[\frac{\Delta\omega f(t)}{\alpha}\right]^3 \pm \cdots\right\}. \end{aligned} \qquad (12.4\text{–}8)$$

Now, if $2\Delta\omega/3\alpha \leq 0.04$ or, equivalently, $\mathrm{BW} \geq 100\,\Delta\omega/3$, then the third term of Eq. (12.4–8) is less than 1 % of the second term for all $f(t)$, and thus $v_o(t)$ varies reasonably linearly with $f(t)$. Specifically, if the bandwidth is chosen at its minimum value

BW = 100 Δω/3, then $v_o(t)$ is given by

$$v_o(t) = I_1 R\sqrt{\tfrac{2}{3}}\left[1 \pm \frac{f(t)}{25\sqrt{2}}\right]. \tag{12.4–9}$$

As we saw in Section 11.3 the quasi-static approximation is valid for the single-tuned circuit provided that (cf. Eq. 11.3–17)

$$\frac{1}{\beta}\left(\frac{\Delta\omega}{\alpha}\right)^2 \ll 1.$$

For the circuit just considered, $(\Delta\omega/\alpha)^2 = 0.0036$; hence, for any value of β greater than 0.36,

$$\frac{1}{\beta}\left(\frac{\Delta\omega}{\alpha}\right)^2 < 0.01,$$

the validity of quasi-static approximation is ensured, and thus Eq. (12.4–9) provides an accurate expression for $v_o(t)$.

Once ω_0, $\Delta\omega$, and ω_m for the input FM signal are known, the parameters of the tuned circuit can be selected. For example, if $\omega_0 = 2\pi(10.7\text{ MHz})$, $\Delta\omega = 2\pi(75\text{ kHz})$, and $\omega_m = 2\pi(15\text{ kHz})$ (commercial FM), then $\text{BW} = 100\,\Delta\omega/3 = 2\pi(1.5\text{ MHz})$ and

$$\omega_c = \omega_0 + \frac{\alpha}{\sqrt{2}} = \omega_0 + \frac{\text{BW}}{2\sqrt{2}} = 2\pi(11.23\text{ MHz}),$$

with the result that $Q_T = \omega_c/\text{BW} = 7.5$. If, in addition, $R = 10\text{ k}\Omega$ and $I_1 = 1\text{ mA}$, then $C = 1/(\text{BW})R = 6.7\text{ pF}$, $L = 1/\omega_c^2 C = 12\,\mu\text{H}$, and

$$v_o(t) = (8.15\text{ V}) + (0.23\text{ V})f(t)$$

($\beta = \Delta\omega/\omega_m = 5$ ensures the validity of the quasi-static approximation).

Balanced Slope Demodulator

As was pointed out in Section 12.3, single-ended demodulators are virtually never used if $\Delta\omega/\omega_0$ is small, because of their high sensitivity to envelope variations on the input signal. Consequently, to make the slope demodulator less sensitive to envelope variations at its input, a balancing path must be introduced. A balanced slope demodulator which not only produces zero output for $\omega_i = \omega_0$ but also has an increased linear operating range is shown in Fig. 12.4–3. For this circuit, balancing is achieved by taking the difference of the output voltages of two single-ended slope demodulators, one tuned above ω_0 and the other tuned below ω_0 by a similar amount.

If again we assume that the envelope detectors do not load the tuned circuits, that the "quasi-static" approximation is valid for determining the voltage across each tuned circuit, and that $i_i(t)$ is given by

$$i_i(t) = I_1 \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\,d\theta\right], \tag{12.4–10}$$

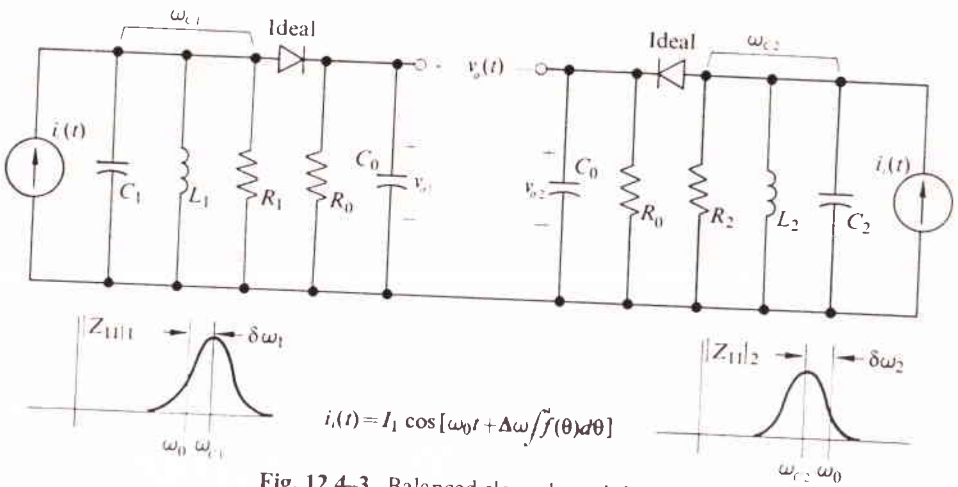


Fig. 12.4–3 Balanced slope demodulator.

then for the demodulator shown in Fig. 12.4–3

$$\begin{aligned} v_o(t) &= v_{o1}(t) - v_{o2}(t) \\ &= I_1\{|Z_{11}[j\omega_0 + j\,\Delta\omega f(t)]|_1 - |Z_{11}[j\omega + j\,\Delta\omega f(t)]|_2\} \\ &= I_1|Z_T[j\omega_0 + j\,\Delta\omega f(t)]|, \end{aligned} \tag{12.4–11}$$

where $|Z_T(j\omega)| = |Z_{11}(j\omega)|_1 - |Z_{11}(j\omega)|_2$ and $Z_{11}(p)_1$ and $Z_{11}(p)_2$ are the input impedances of the parallel $R_1L_1C_1$ and $R_2L_2C_2$ circuits respectively. The transfer characteristic now relating $\omega_i(t)$ and $v_o(t)$ is shown in Fig. 12.4–4.

From Fig. 12.4–4 the reason for the increased linearity becomes apparent. If the two individual characteristics shown in Fig. 12.4–4 are correctly adjusted, they combine in such a way as to form a $v_o(t)$-vs.-$\omega_i(t)$ characteristic which has odd

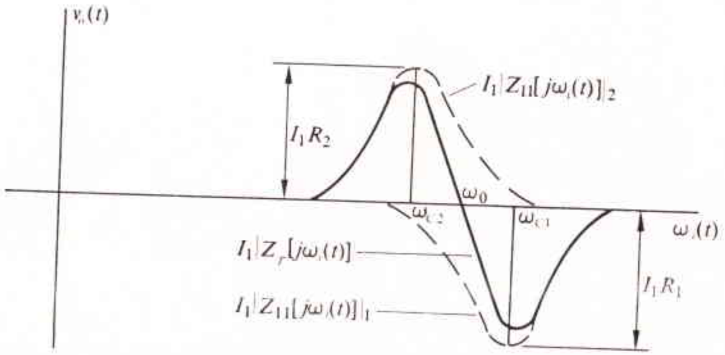


Fig. 12.4–4 Transfer characteristic relating $\omega_i(t)$ and $v_o(t)$ for the circuit of Fig. 12.4–3.

symmetry about ω_0. Thus if $|Z_T(j\omega)|$ is expanded in a Taylor series in ω about ω_0, the even terms vanish:

$$|Z_T(j\omega)| = |Z_T(j\omega_0)|'(\omega - \omega_0) + |Z_T(j\omega_0)|'''\frac{(\omega - \omega_0)^3}{3!}$$
$$+ |Z_T(j\omega_0)|^{\mathrm{v}}\frac{(\omega - \omega_0)^5}{5!} + \cdots. \qquad (12.4\text{–}12)$$

Therefore, ω_{c1} and ω_{c2} can now be chosen such that $|Z_T(j\omega_0)|''' = 0$ [since $|Z_T(j\omega_0)|''$ is already zero], with the result that the first term introducing a nonlinearity in $|Z_T(j\omega)|$, and in turn $v_o(t)$, involves the fifth derivative of $|Z_T(j\omega)|\,|_{\omega=\omega_0}$.

If Q_T for both the tuned circuits of Fig. 12.4–3 is greater than 10, then $|Z_T(j\omega)|$ may be closely approximated by

$$|Z_T(j\omega)| = \frac{R_1}{\sqrt{1 + \left(\dfrac{\omega - \omega_{c1}}{\alpha_1}\right)^2}} - \frac{R_2}{\sqrt{1 + \left(\dfrac{\omega - \omega_{c2}}{\alpha_2}\right)^2}}, \qquad (12.4\text{–}13)$$

where $\alpha_1 = 1/2R_1C_1 = \mathrm{BW}_1/2$ and $\alpha_2 = 1/2R_2C_2 = \mathrm{BW}_2/2$. For this case, odd symmetry in $|Z_T(j\omega)|$ about ω_0 is obtained by choosing $R_1 = R_2 = R$, $\alpha_1 = \alpha_2 = \alpha = \mathrm{BW}/2$, $\omega_{c1} = \omega_0 + \delta\omega$, and $\omega_{c2} = \omega_0 - \delta\omega$. In addition (as the reader should verify), $|Z_T(j\omega_0)|''' = 0$ if $\delta\omega$ is chosen as

$$\delta\omega = \sqrt{\frac{3}{2}}\alpha = \sqrt{\frac{3}{2}}\frac{\mathrm{BW}}{2}. \qquad (12.4\text{–}14)$$

With this choice for $\delta\omega$, the Taylor series expansion for $|Z_T(j\omega)|$ reduces to

$$|Z_T(j\omega)| = \frac{4R}{5}\sqrt{\frac{3}{5}}\left[\frac{\omega - \omega_0}{\alpha} - \frac{54}{625}\left(\frac{\omega - \omega_0}{\alpha}\right)^5 + \cdots\right] \qquad (12.4\text{–}15)$$

and in turn

$$v_o(t) = \frac{4I_1R}{5}\sqrt{\frac{3}{5}}\left[\frac{\Delta\omega f(t)}{\alpha} - \frac{54}{625}\left(\frac{\Delta\omega f(t)}{\alpha}\right)^5 + \cdots\right]. \qquad (12.4\text{–}16)$$

For this case, if $\Delta\omega/\alpha \le 0.585$, then the second term of Eq. (12.4–16) is less than 1% of the first term for all values of $f(t)$; thus $v_o(t)$ varies in direct proportion to $f(t)$.

However, a value of $\Delta\omega/\alpha$ as large as 0.585 puts a strain on the quasi-static approximation for values of β less than 7 (cf. Eq. 11.3–17). Therefore, $\Delta\omega/\alpha$ is usually chosen to be $\frac{1}{2}$, a value which ensures the validity of Eq. (11.3–17) with β as low as 5

With $\Delta\omega/\alpha = \frac{1}{2}$ (or $\mathrm{BW} = 4\,\Delta\omega$), $v_o(t)$ given by Eq. (12.4–16) simplifies to

$$v_o(t) = \frac{2I_1R}{5}\sqrt{\tfrac{3}{5}}f(t). \qquad (12.4\text{–}17)$$

Hence, if $I_1 = 1$ mA and $R = 10\,\mathrm{k}\Omega$, $v_o(t) = (3.1\ \mathrm{V})f(t)$; thus the balanced slope demodulator, in addition to its other advantages, provides a coefficient for the $f(t)$-term in its output which is more than 13 times as large as the corresponding coefficient for the single-ended slope demodulator.

Here again, parameters for the tuned circuits of the balanced slope demodulator may be chosen once ω_0, $\Delta\omega$, and ω_m are specified. Using the commercial FM values, we obtain $\mathrm{BW} = 4\,\Delta\omega = 2\pi(300\ \mathrm{kHz})$, $\delta\omega = 2\pi(184\ \mathrm{kHz})$, $\omega_{c1} = 2\pi(10.884\ \mathrm{MHz})$, and $\omega_{c2} = 2\pi(10.516\ \mathrm{MHz})$; hence $Q_{T1} = 36.3$, $Q_{T2} = 35.2$, $C_1 = C_2 = 53$ pF, $L_1 = 4.04\ \mu$H, and $L_2 = 4.55\ \mu$H.

A possible physical circuit for realizing the balanced slope demodulator is shown in Fig. 12.4–5. For this circuit the first differential pair not only provides the two identical current drives $i_i(t)$, but also acts as a limiter. (Actually the currents at the collectors of Q_1 and Q_2 are 180° out of phase with each other. The envelope detectors, however, are not sensitive to this phase inversion.) The diodes in the detectors are

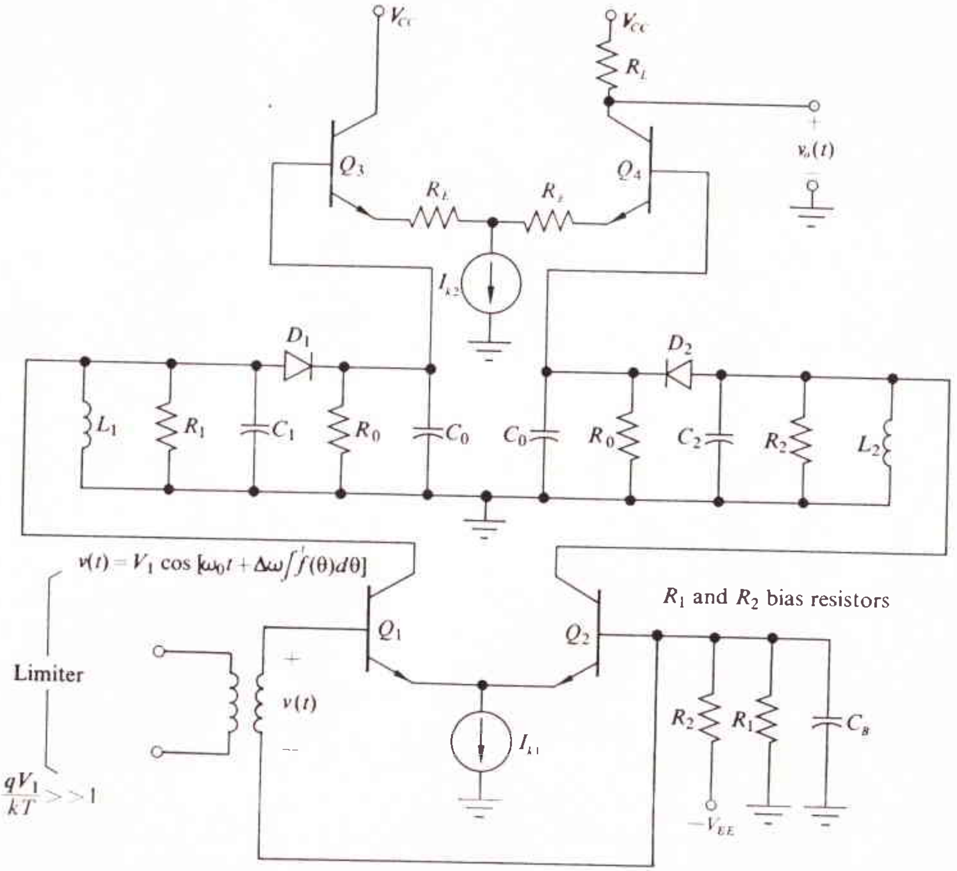


Fig. 12.4–5 Practical embodiment of balanced slope demodulator.

ade of germanium to minimize their effective V_0; however, if the diodes are matched, e V_0-term is balanced out at the output. Finally, the second differential pair accepts e differential input and provides a single-ended output. The resistors are added in ries with the emitters of this differential pair to increase the linear operating range nd reduce the loading of the differential pair on the envelope detectors. The modu- ted input to this circuit is the transformer-coupled voltage $v(t)$ appearing at the ase of Q_1.

nvelope Detector Loading

n many applications it is not desirable to make the output resistance R_0 of the nvelope detector in the slope demodulator so large that it does not load the tuned ircuit. However, provided that C_0 looks like an open circuit in comparison with $_0$ over the band of frequencies occupied by $f(t)$, that is, $0 \leq \omega \leq \omega_m$, the loading of e envelope detector is readily accounted for. In particular, with C_0 an open circuit $f(t)$, the envelope detector functions in a static fashion and its primary effect is to wer the Q_T of the tuned circuit by reflecting a resistor of value $R_0/2$ in parallel with e parallel RLC circuit. For example, if the tuned circuit has a parallel resistance of 0 kΩ and $R_0 = 40$ kΩ, the loading of the envelope detector may be accounted for by ecreasing the total parallel resistance $R_T = R\|(R_0/2)$ to 10 kΩ. This value of R_T then used in all the expressions previously derived.

eries-Tuned Balanced Slope Demodulator

ries resonant circuits can also be used to obtain a slope demodulator. For this dual rangement a common voltage must drive the two series-tuned circuits, and the ifference of the envelopes of their currents must be used to provide an output. A rcuit for accomplishing this is illustrated in Fig. 12.4–6. For this demodulator an verage envelope detector is used instead of the peak envelope detector; thus the strictions on D and β of the input FM signal are considerably relaxed. Specificall and β must determine a point under the curve of Fig. 12.2–7 instead of determinin point under the more restrictive curve of Fig. 12.2–9.

For the circuit of Fig. 12.4–6, if $V_0 \approx 0$ for the diodes and transistors, then $v(t)$ pears directly across the series-tuned circuits, with the result that the envelope of (t) is given by

$$V_1|Y_{11}[j\omega_0 + j\Delta\omega f(t)]|_1$$

d the envelope of $i_2(t)$ is given by

$$V_1|Y_{11}[j\omega_0 + j\,\Delta\omega f(t)]|_2,$$

here $Y_{11}(p)_1$ and $Y_{11}(p)_2$ are the admittances of the series $r_1L_1C_1$ and r_2L_2C sonant circuits respectively. For each tuned circuit, the series resistance includ y input resistance of the transistor-diode combinations. Consequently, if the ban

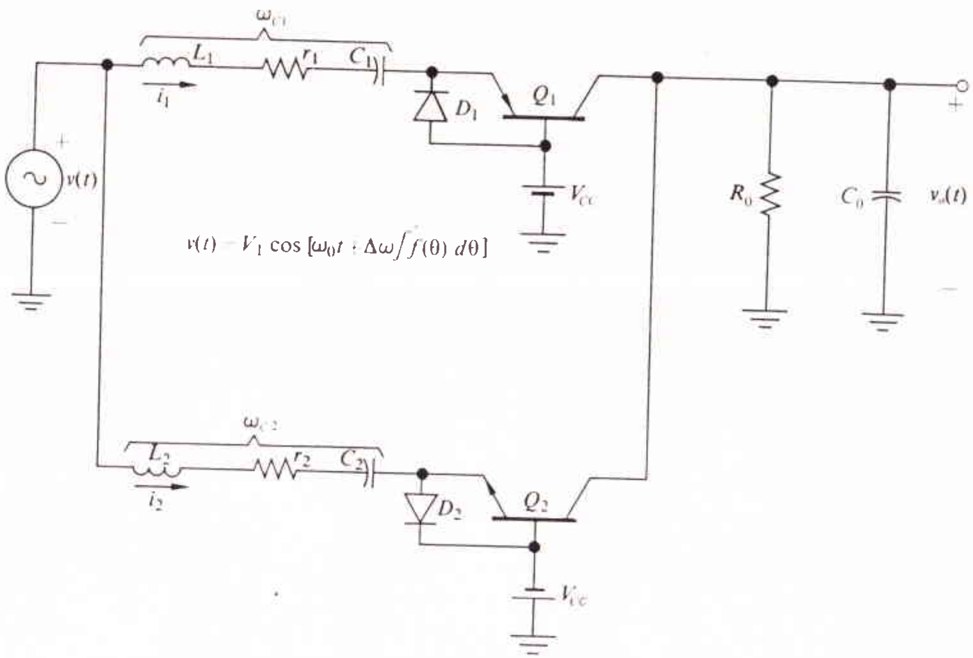


Fig. 12.4–6 Series-tuned balanced slope demodulator.

width of the R_0–C_0 low-pass filter is greater than ω_m, then $v_o(t)$ is given by

$$\begin{aligned} v_o(t) &= \frac{\alpha_T V_1 R_0}{\pi}\{|Y_{11}[j\omega_i(t)]|_1 - |Y_{11}[j\omega_i(t)]|_2\} \\ &= \frac{\alpha_T V_1 R_0}{\pi}|Y_T[j\omega_i(t)]|, \end{aligned} \tag{12.4–18}$$

where $|Y_T(j\omega)| = |Y_{11}(j\omega)| = |Y_{11}(j\omega)|_1 - |Y_{11}(j\omega)|_2$. (A subscript T is used to differentiate the transistor α_T from the α of the tuned circuit.)

It is apparent that Eq. (12.4–18) is of the same form as Eq. (12.4–11); hence the same considerations employed in choosing parameter values for the tuned circuits of Fig. 12.4–3 apply to the tuned circuits of Fig. 12.4–6. In particular, if $Q_T > 10$ for both of the tuned circuits of Fig. 12.4–6, then

$$v_o(t) = \frac{\alpha_T V_1 R_0}{\pi}\left\{\frac{1}{r_1\sqrt{1 + \left[\dfrac{\omega_i(t) - \omega_{c1}}{\alpha_1}\right]^2}} - \frac{1}{r_2\sqrt{1 + \left[\dfrac{\omega_i(t) - \omega_{c2}}{\alpha_2}\right]^2}}\right\}, \tag{12.4–19}$$

where

$$\omega_{c1} = \frac{1}{\sqrt{L_1C_1}}, \qquad \omega_{c2} = \frac{1}{\sqrt{L_1C_2}},$$

$$\alpha_1 = \frac{r_1}{2L_1}, \qquad \text{and} \qquad \alpha_2 = \frac{r_2}{2L_2}.$$

nce, if we again choose $\alpha_1 = \alpha_2 = \alpha, r_1 = r_2 = r, \omega_{c1} = \omega_0 + \delta\omega, \omega_{c2} = \omega_0 - \delta\omega$, d $\delta\omega = \sqrt{\frac{3}{2}}\alpha$, then provided that $\Delta\omega/\alpha$ is kept less than $\frac{1}{2}$, $v_o(t)$ reduces to

$$v_o(t) = \frac{4\alpha_T V_1 R_0}{5\pi r} \sqrt{\frac{3}{5}} \frac{\Delta\omega f(t)}{\alpha}. \tag{12.4–20}$$

As with the Clarke-Hess demodulator discussed in Section 12.3, R_0 should be osen sufficiently large to maximize $v_o(t)$, but not so large that either transistor in the cuit of Fig. 12.4–6 saturates on the peaks of $f(t)$.

Although we have assumed $V_0 \approx 0$ in this analysis, the expression for $v_o(t)$ ven by Eq. (12.4–20) is not appreciably affected if V_0 is a few tenths of a volt. This sensitivity is a direct result of the balanced nature of the circuit: thus, to aid in this lancing, every effort should be made to match the input characteristics of the transtors and diodes of Fig. 12.4–6. If this is not done, some variations in center freency with input amplitude will result.

rystal Slope Demodulator

place of the two tuned circuits of Fig. 12.4–6, two crystal resonators may be inserted, ne with its series resonance slightly above ω_0 and the other with its series resonance ightly below ω_0. The result is a slope demodulator with an extremely stable center equency and an output voltage which is extremely sensitive to small deviations out ω_0.

Smith† has reported such a discriminator in which, instead of separate crystals, ree sections of a single quartz crystal are mechanically coupled. He presents results r a detector with a center frequency of 15 MHz and a linear bandwidth of ± 3 kHz. is device is reported to be generally useful in the 5–40 MHz range and to have posible useful bandwidths of up to 0.2% of the center frequency.

A single crystal used with a balancing branch is often employed as the disriminator in the frequency stabilizing situation described in Section 11.9. In such ituations it is normally more satisfactory to operate the crystal in the vicinity of the ero in its impedance function rather than near its parallel resonant frequency. (Secion 6.7 discusses the crystal's impedance function in some detail.)

2.5 TIME-DELAY DIFFERENTIATOR, TIME-DELAY EMODULATOR, FOSTER-SEELEY DEMODULATOR, AND RATIO DETECTOR

time-delay differentiator is a linear network which implements the fundamental efinition of a derivative,

$$\frac{dv(t)}{dt} = \lim_{t_0 \to 0} \frac{v(t) - v(t - t_0)}{t_0}. \tag{12.5–1}$$

he block diagram for such a network is shown in Fig. 12.5–1. In this diagram we

---

† Warren L. Smith, "The Monolithic FM Discriminator: A New Piezoelectric Device," *IEE Trans. on Communication Technology*, **Com–16**, No. 3, pp. 460–463 (June 1968).

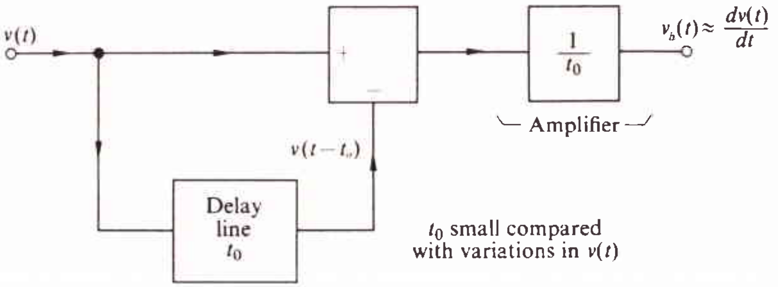

v(t)
+
−
$\frac{1}{t_0}$
$v_b(t) \approx \frac{dv(t)}{dt}$
Amplifier
v(t − t₀)
Delay line t_0
t_0 small compared with variations in v(t)

Fig. 12.5–1 Block diagram of a time-delay differentiator.

observe that a delayed version of $v(t)$ is subtracted from $v(t)$ and the resultant signal is amplified by $1/t_0$. In general, t_0 must be chosen equal to zero to realize the exact derivative of $v(t)$; however, as a practical matter, so long as t_0 is small in comparison with the time interval during which variations in $v(t)$ occur, the derivative of $v(t)$ is approximated quite accurately.

It is apparent that t_0 should not be chosen any smaller than necessary, since such a choice requires an additional amplifier gain $(1/t_0)$ to bring the signal up to a useful level. Therefore, we shall now determine the maximum value of t_0 which permits the network to function as a differentiator (i.e., to convert variations in instantaneous frequency into envelope variations) for the case where $v(t)$ has the form

$$v(t) = A \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\, d\theta\right]. \tag{12.5–2}$$

In the next section we shall again determine the maximum permissible size of t_0 where $v(t)$ is a square-wave FM signal given by Eq. (12.2–1b). Once the maximum size of t_0 is determined, we shall implement the block diagram of Fig. 12.5–1 with practical circuits and then interconnect them with the various amplitude demodulators to form frequency demodulators.

For the purpose of analysis we shall consider the more general time-delay differentiator, shown in Fig. 12.5–2, from which the amplifier has been omitted and the

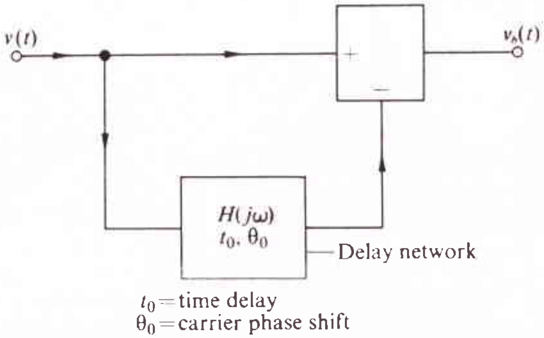


Fig. 12.5–2 General time-delay differentiator.

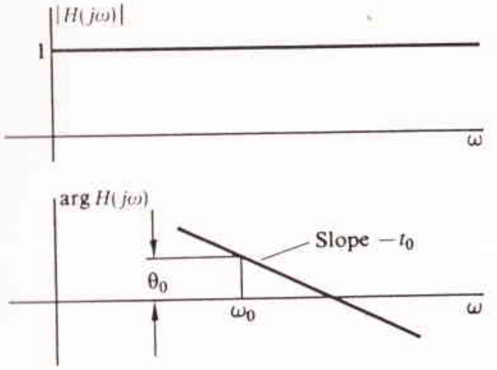


Fig. 12.5–3 Magnitude and phase characteristic of the delay network of Fig. 12.5–2.

vision for a phase shift of θ_0 at ω_0 in the delay network has been included. Figure .5–3 illustrates the magnitude and phase characteristic of the delay network of Fig. .5–2. Note that, if $\theta_0 = -\omega_0 t_0$, then the phase characteristic passes through the igin and the normal delay line shown in Fig. 12.5–1 results.

If we apply $v(t)$ given by Eq. (12.5–2) to the time-delay differentiator shown in g. 12.5–2, then with the aid of Eq. (11.3–3) we can write

$$
\begin{aligned}
v_b(t) &= A\cos\left[\omega_0 t + \Delta\omega\int^{t} f(\theta)\,d\theta\right] - A\cos\left[\omega_0 t + \Delta\omega\int^{t-t_0} f(\theta)\,d\theta + \theta_0\right] \\
&= \overbrace{-2A\sin\tfrac{1}{2}\left[\Delta\omega\int_{t-t_0}^{t} f(\theta)\,d\theta - \theta_0\right]}^{\text{Envelope term: } g(t)} \\
&\quad \underbrace{\times \sin\left\{\omega_0 t + \Delta\omega\int^{t} f(\theta)\,d\theta - \tfrac{1}{2}\left[\Delta\omega\int_{t-t_0}^{t} f(\theta)\,d\theta - \theta_0\right]\right\}}_{\text{High-frequency term}}.
\end{aligned}
\tag{12.5–3}
$$

e second expression in Eq. (12.5–3) is obtained with the aid of the trigonometric entity

$$\cos x - \cos y = -2\sin\frac{x+y}{2}\sin\frac{x-y}{2}.$$

nce $v_b(t)$ will be placed through an envelope detector, only the envelope $g(t)$ of $v_b(t)$ ed be considered in determining the maximum permissible value for t_0. More ecifically, the maximum value of t_0 is that value which still permits $g(t)$ given by

$$g(t) = 2A\sin\tfrac{1}{2}\left[\Delta\omega\int_{t-t_0}^{t} f(\theta)\,d\theta - \theta_0\right] \tag{12.5–4}$$

vary in direct proportion to $f(t)$.

The envelope $g(t)$ may be arranged in a form in which its dependence on t_0 is more explicit by defining

$$
\begin{aligned}
k(t) &= \frac{\Delta\omega}{2}\int_{t-t_0}^{t} f(\theta)\,d\theta \\
&= \frac{\Delta\omega}{2}\int^{t} f(\theta)\,d\theta - \frac{\Delta\omega}{2}\int^{t-t_0} f(\theta)\,d\theta.
\end{aligned}
\tag{12.5–5}
$$

Now by taking the Fourier transform of Eq. (12.5–5) we obtain

$$
\begin{aligned}
K(\omega) &= \frac{\Delta\omega F(\omega)}{2}\,\frac{1}{j\omega}(1 - e^{-j\omega t_0}) \\
&= F(\omega)e^{-j\omega t_0/2}\,\frac{\Delta\omega t_0}{2}\,\frac{\sin(\omega t_0/2)}{\omega t_0/2} \\
&= F(\omega)H_D(j\omega),
\end{aligned}
\tag{12.5–6}
$$

where $K(\omega)$ and $F(\omega)$ are the Fourier transforms of $k(t)$ and $f(t)$ respectively and

$$H_D(j\omega) = e^{-j\omega t_0/2}\,\frac{\Delta\omega t_0}{2}\,\frac{\sin(\omega t_0/2)}{\omega t_0/2}.$$

A plot of $|H_D(j\omega)|$ and $\arg H_D(j\omega)$ vs. ω is shown in Fig. 12.5–4, from which we observe that $|H_D(j\omega)|$ is essentially constant at $\Delta\omega t_0/2$ for $\omega < 2/t_0$; hence, if $\omega_m < 2/t_0$, where ω_m is the limit on the spectrum of $F(\omega)$, then

$$K(\omega) = F(\omega)H_D(\omega) \approx \frac{\Delta\omega t_0}{2}e^{-j\omega t_0/2}F(\omega), \tag{12.5–7}$$

or equivalently,

$$k(t) = \frac{\Delta\omega t_0}{2}f\left(t - \frac{t_0}{2}\right). \tag{12.5–8}$$

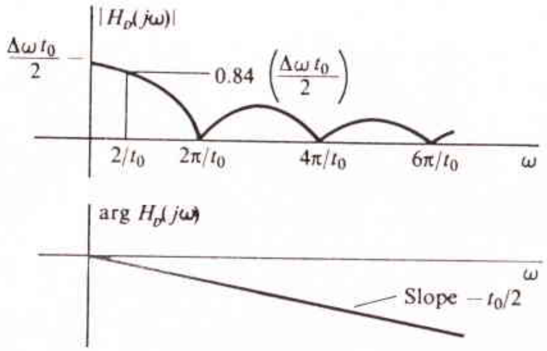


Fig. 12.5–4 Plot of the magnitude and phase characteristic of $H_D(j\omega)$.

Thus we see that $k(t)$ is proportional to a delayed version† of $f(t)$ provided that

$$t_0 < \frac{2}{\omega_m}. \tag{12.5–9}$$

With $k(t)$ given by Eq. (12.5–8), $g(t)$ simplifies to

$$g(t) = 2A \sin\left[\frac{\Delta\omega t_0}{2} f\left(t - \frac{t_0}{2}\right) - \frac{\theta_0}{2}\right]. \tag{12.5–10}$$

The transfer characteristic relating $g(t)$ and $(\Delta\omega t_0/2)\, f(t - t_0/2)$ is shown in Fig. 12.5–5, from which we observe that only if

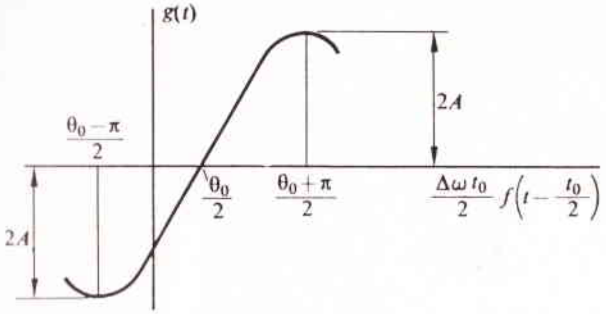


Fig. 12.5–5 Transfer characteristic relating $g(t)$ and $(\Delta\omega t_0/2) f(t - t_0/2)$.

$$\frac{\Delta\omega t_0}{2} f\left(t - \frac{t_0}{2}\right) \ll 1 \qquad \text{for all } t$$

does $g(t)$ vary in proportion to $f(t - t_0/2)$ as desired. More specifically, if we expand $g(t)$ given by Eq. (12.5–10) in the form

$$g(t) = 2A \sin k(t) \cos\frac{\theta_0}{2} - 2A \cos k(t) \sin\frac{\theta_0}{2}, \tag{12.5–11}$$

we observe that, if $k(t)$ is kept less than 0.2 for all t, then $\sin k(t) \approx k(t)$ and $\cos k(t) \approx 1$ (within 2%), with the result that

$$\begin{aligned} g(t) &= 2Ak(t)\cos\frac{\theta_0}{2} - 2A\sin\frac{\theta_0}{2} \\ &= A\,\Delta\omega t_0 f\left(t - \frac{t_0}{2}\right)\cos\frac{\theta_0}{2} - 2A\sin\frac{\theta_0}{2}. \end{aligned} \tag{12.5–12}$$

The requirement that $k(t) < 0.2$ is equivalent to

$$\Delta\omega t_0 < 0.4, \tag{12.5–13}$$

as is readily seen from Eq. (12.5–8).

---

† A slight delay is unimportant in most communication applications.

The desired proportionality between $g(t)$ and $f(t - t_0/2)$, as expressed in Eq. (12.5–12), requires that both Eqs. (12.5–9) and (12.5–13) be satisfied. When $\beta > 0.4/2 = 0.2$ (which is usually the case in practice), Eq. (12.5–13) is more restrictive than Eq. (12.5–9) hence for this case the maximum size of t_0 is restricted to $0.4/\Delta\omega$.

In general, a value for θ_0 must also be selected. It is apparent that, if θ_0 is chosen as 0, 2π, 4π, etc., the component of $g(t)$ proportional to $f(t_0/2)$ is maximized. The difficulty with this choice is that $g(t)$ is not positive (or negative) for all t and thus peak or average envelope detection is impossible. To overcome this difficulty, θ_0 must be chosen such that (cf. Eq. 12.5–12)

$$\Delta\omega t_0 \cos\frac{\theta_0}{2} \le 2\sin\frac{\theta_0}{2},$$

or equivalently,

$$\tan\frac{\theta_0}{2} \ge \frac{\Delta\omega t_0}{2}. \tag{12.5–14}$$

Since $\Delta\omega t_0 < 0.4$, Eq. (12.2–33) is satisfied provided that

$$\theta_0 > 0.4(22.7°). \tag{12.5–15}$$

In many practical circuits a phase shift at the carrier frequency of $\theta_0 = \pm\pi/2$ is readily obtained; thus this value is usually employed, with the result that Eq. (12.5–12) simplifies to

$$g(t) = \mp A\sqrt{2}\left[1 \mp \frac{\Delta\omega t_0}{2} f\left(t - \frac{t_0}{2}\right)\right]. \tag{12.5–16}$$

Consequently, if the time-delay differentiator (with $\theta_0 = \pm\pi/2$) is followed by an envelope detector as shown in Fig. 12.5–6 and if t_0 is chosen at its maximum value of $0.4/\Delta\omega$, then the resultant frequency demodulator output is given by

$$v_o(t) = AK_M\sqrt{2}[1 \mp 0.2f(t - t_0/2)], \tag{12.5–17}$$

where K_M is the amplitude demodulator constant and the − and + correspond to $\theta_0 = \pi/2$ and $\theta_0 = -\pi/2$ respectively.

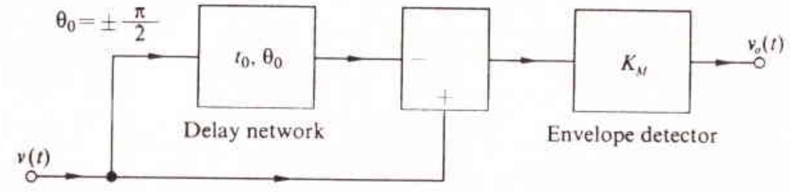


Fig. 12.5–6 Single-ended time-delay frequency demodulator.

The single-ended time-delay demodulator finds wide application in demodulating FM microwave signals. Two paths are provided for $v(t)$ through waveguides or transmission lines of differing lengths. The difference in length is adjusted so that one signal experiences a time delay t_0 and a phase shift $\theta_0 = -\omega_0 t_0 = -3\pi/2 = -\pi/2 - \pi$ (that is, a phase shift of $-\pi/2$ plus an inversion). The output of the two waveguides (or transmission lines) is then combined and envelope-detected with a crystal detector inserted in the guide. The detector output is the demodulated FM signal. A similar scheme should apply to the demodulation of frequency-modulated laser beams.

A balanced frequency demodulator employing a time-delay differentiator is shown in block diagram form in Fig. 12.5–7. In this diagram two single-ended time-delay frequency demodulators, one with $\theta_0 = -\pi/2$ and the other with $\theta_0 = \pi/2$, are arranged in such a way that their outputs subtract, with the result that $v_o(t)$ is given by

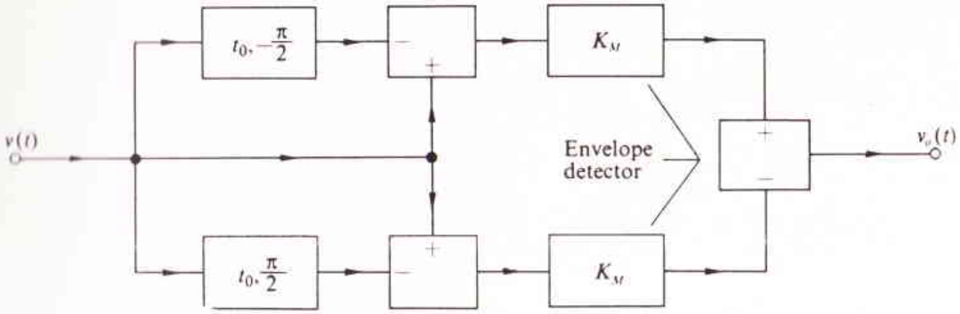


Fig. 12.5–7 Balanced time-delay frequency demodulator.

$$v_o(t) = \sqrt{2}AK_M\,\Delta\omega t_0 f(t - t_0/2). \tag{12.5–18}$$

For this case, not only are all the advantages of balancing obtained, but also the coefficient of the $f(t - t_0/2)$-term is increased by a factor of 2 over the single-ended demodulator.

By noting that a phase shift of $\pi/2$ can be obtained through a phase shift of $-\pi/2$ plus a signal inversion (a phase shift of $-\pi$), we may rearrange the balanced time-delay demodulator of Fig. 12.5–7 in a form which requires only phase shifts of $-\pi/2$ at the carrier frequency. (A phase shift of $-\pi/2$ is more readily realized in a physical circuit than is a phase shift of $+\pi/2$.) Two versions of the resultant circuit are shown in Fig. 12.5–8.

When an actual piece of delay line is used to implement the circuit, the single-delay-line circuit is the obvious choice. For this case, $\theta_0 = -\omega_0 t_0$; hence, for $\theta_0 = -\pi/2$,

$$t_0 = T_0/4,$$

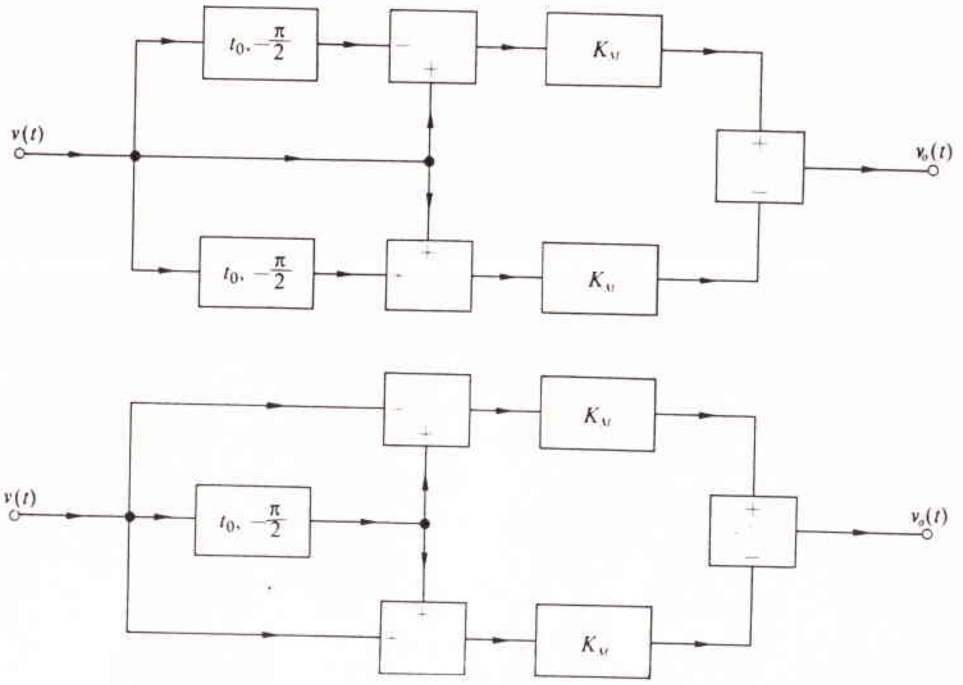


Fig. 12.5–8 Alternative forms for the balanced time-delay demodulator.

where T_0 is the carrier period. Thus to construct a delay-line demodulator at $T_0 = 100$ nsec ($f_0 = 10$ MHz), one requires a minimum overall delay of 25 nsec. Since an ordinary coaxial line has a delay of the order of 5 nsec/m, such a detector will be rather cumbersome if built with this type of line. Fortunately helically wound delay lines exist† with delays of up to 30 nsec/cm or 3 μsec/10 m. Thus, with the use of distributed lines, delay-line demodulators are a practical possibility for values of f_0 as low as 800 kHz.

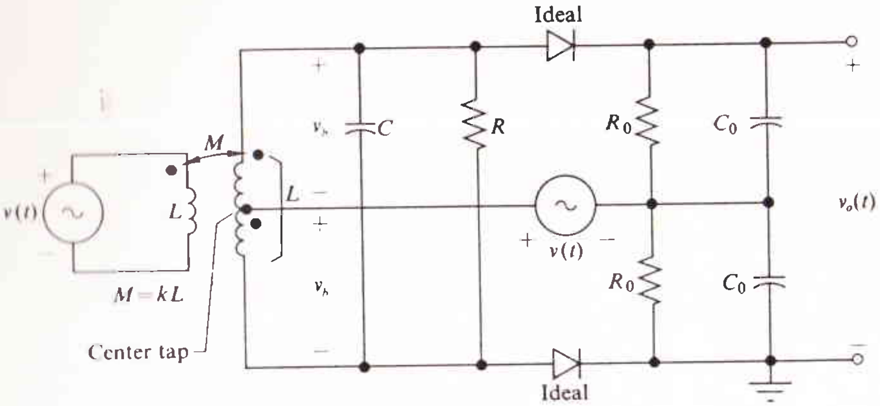
The Foster–Seeley Demodulator

Many existing FM demodulators use the nearly constant phase characteristic of a tuned circuit near resonance to approximate the delay line and thus to realize a delay-line type of demodulator. In the analysis of such circuits, one should remember that two layers of approximations are being used. In the first layer, one is approximating an ideal differentiator with a delay line; in the second layer, one is approximating the delay line by the tuned circuit. In the second approximation, one normally further neglects the circuit-induced amplitude variations.

---

† See J. Millman and H. Taub, *Pulse and Digital Circuits*, McGraw-Hill, New York (1956). Chapter 10, Section 1 describes a number of lines with delays from 140 nsec/m to 30,000 nsec/m.

Figure 12.5–9 illustrates the widely used Foster-Seeley FM demodulator.† As e shall show, in this circuit the tuned transformer introduces a time delay t_0 and a hase shift of $-\pi/2$ between $v(t)$ and $v_b(t)$. The voltage $v_b(t)$ is then combined with $v(t)$ a fashion identical with that of the block diagram of Fig. 12.5–8 to yield $v_o(t)$ given y Eq. (12.5–18). Since peak envelope detection is employed, $K_M = 1$.



ig. 12.5–9 Foster-Seeley demodulator which physically implements the block diagram of ig. 12.5–8 (top).

To obtain a suitable expression for t_0 in terms of the parameters of the tuned ircuit of Fig. 12.5–9, let us assume that the envelope detectors do not load the tuned ransformer circuit. (If they do, as was pointed out in the previous section, a resistor $_0$ must be placed in parallel with R.) Neglecting detector loading, we can determine $_b(t)$ from the model of the tuned transformer illustrated in Fig. 12.5–10 (cf. Fig. A–3c).

With this model, $v_b(t)$ is related to $v(t)$ by the transfer function

$$H(p) = \frac{V_b(p)}{V(p)} = \frac{k\omega_0^2/2}{p^2 + 2\alpha p + \omega_0^2}, \tag{12.5–19}$$

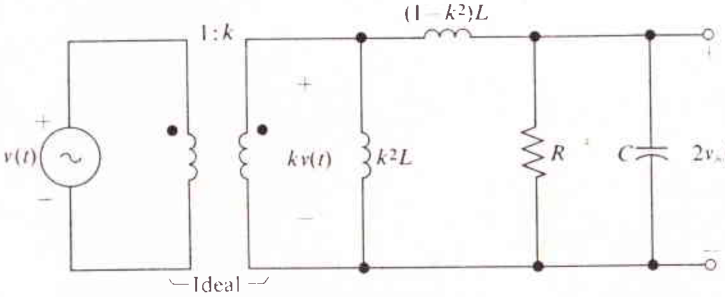


Fig. 12.5–10 Model for the tuned transformer of Fig. 12.5–9.

D. F. Foster and S. H. Seeley, *Proc. IRE*, **25**, p. 289 (1937).

---

where $\omega_0 = 1/\sqrt{(1-k^2)LC}$ and $\alpha = 1/2RC$. The pole-zero diagram of $H(p)$ is shown in Fig. 12.5–11. Note, however, that this pole-zero diagram is identical with that of Fig. 3.1–5 (except for the scale factor), for which θ_0 was found to equal $-\pi/2$. In addition, for $Q_T = \omega_0 RC > 10$ (cf. Eq. 3.1–9), $H(j\omega)$ may be closely approximated as

$$H(j\omega) = \frac{(kQ_T/2)e^{-j(\pi/2)}}{1 + j[(\omega - \omega_0)/\alpha]}$$
$$= \frac{kQ_T e^{-j\{\tan^{-1}[(\omega-\omega_0)/\alpha] + \pi/2\}}}{2\sqrt{1 + [(\omega - \omega_0)/\alpha]^2}}. \tag{12.5–20}$$

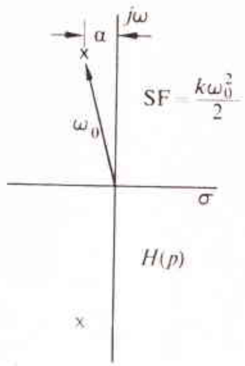


Fig. 12.5–11 Pole-zero diagram of $H(p)$.

For $|(\omega - \omega_0)/\alpha| < 0.2$, $\tan^{-1}[(\omega - \omega_0)/\alpha]$ may be approximated by $(\omega - \omega_0)/\alpha$ within 2% and $\sqrt{1 + [(\omega - \omega_0)/\alpha]^2}$ may be approximated by unity within 2%. Thus, if we operate sufficiently close to ω_0, $H(j\omega)$ simplifies to

$$H(j\omega) = \frac{kQ_T}{2} e^{-j[(\omega-\omega_0)/\alpha + \pi/2]}, \tag{12.5–21}$$

which is the transfer function of a delay network with a delay of $t_0 = 1/\alpha$, with a carrier phase shift of $-\pi/2$, and with a scale factor of $kQ_T/2$. By choosing

$$\frac{kQ_T}{2} = 1, \tag{12.5–22}$$

we obtain the desired scale factor for the delay network.

For proper operation, the bandwidth BW of the FM signal $v(t)$ must lie within the range over which $H(j\omega)$ appears as a delay network: that is,

$$\frac{\text{BW}}{2\alpha} = \frac{(\text{BW})t_0}{2} \le 0.2. \tag{12.5–23}$$

owever, since BW/2 is greater than $\Delta\omega$, Eq. (12.5–23) is more restrictive than Eq. 2.5–13); hence the maximum size of t_0 [and in turn $v_o(t)$] is limited in this case by ie physical implementation of the delay network.

At this point we are in a position to determine the parameters of the tuned ansformer once ω_0, $\Delta\omega$, and ω_m are specified. Consider, for example, $\omega_0 =$ t(10.7 MHz), $\Delta\omega = 2\pi(75\text{ kHz})$, and $\omega_m = 2\pi(15\text{ kHz})$. For this case $\beta = 5$ and, om Fig. 11.1–4, BW = 3.2 $\Delta\omega$; thus from Eq. (12.5–23) we see that the maximum alue of t_0 is

$$t_0 = \frac{1}{\alpha} = \frac{0.4}{\text{BW}} = \frac{0.4}{3.2\,\Delta\omega} = 266\text{ nsec},$$

r $\alpha = 3.76$ Mrad/sec. Since $Q_T = \omega_0/2\alpha = 9$, the coefficient of coupling must be djusted to equal $2/Q_T = 0.222$.

Now if $R = 10\text{ k}\Omega$, then

$$C = \frac{1}{2\alpha R} = 13.3\text{ pF} \qquad \text{and} \qquad L = \frac{1}{\omega_0^2 C(1-k^2)} = 17.6\ \mu\text{H}.$$

n addition, if $A = 5$ V, from Eq. (12.5–18) we obtain

$$v_o(t) = (0.886\text{ V})f'(t - 133\text{ nsec}) \approx (0.886\text{ V})f'(t).$$

n writing the expression for $v_o(t)$ we implicitly assumed that the diodes in the envelope letector were ideal. Even if they are not, so long as their characteristics are matched. he V_0-term they introduce cancels in the output.

A final version for the circuit of Fig. 12.5–9 is shown in Fig. 12.5–11. In this ircuit, $v(t)$ is developed from a current source drive by tuning the input of the transormer to ω_0; that is, $1/\sqrt{LC_1} = \omega_0$. In addition, the "floating" voltage source $v(t)$ s developed across the RF choke by placing a coupling capacitor from the top of the

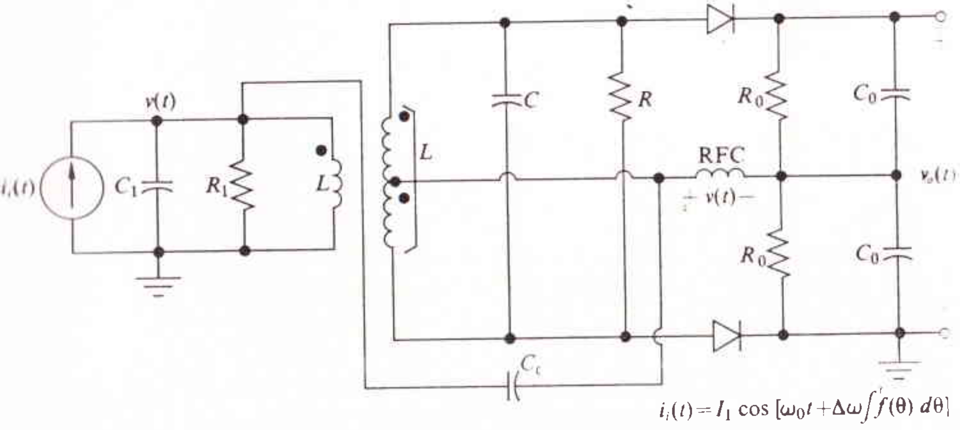


Fig. 12.5–12 Complete Foster–Seeley demodulator.

input-tuned circuit to one end of the RF choke. The other end is grounded to high frequencies through C_0.

If the primary tuned circuit for the circuit of Fig. 12.5–10 has approximately the same Q_T and thus the same bandwidth as the secondary tuned circuit, the bandwidth of $i_i(t)$ is concentrated over the range of frequencies where the input impedance of the tuned transformer appears resistive. Hence

$$A = I_1 R_{\text{in}},$$

where R_{in} is the input resistance of the tuned transformer evaluated at ω_0 and A is the amplitude of $v(t)$. Consequently, once R_{in} is known, $v(t)$ can be determined directly from $i_i(t)$ and the previous analysis may be employed to determine $v_o(t)$. The value for R_{in} is readily shown to be given by

$$R_{\text{in}} = R_1 \| \frac{R}{4}.\dagger$$

Thus if $R = 10\text{ k}\Omega$, $R_1 = 2.5\text{ k}\Omega$, and

$$i_i(t) = (4\text{ mA}) \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\,d\theta\right],$$

then

$$v(t) = (5\text{ V}) \cos\left[\omega_0 t + \Delta\omega \int^t f(\theta)\,d\theta\right].$$

The Ratio Detector

If one of the diodes of the Foster-Seeley demodulator is reversed and a large capacitor is placed in the circuit as shown in Fig. 12.5–13, the ratio detector results.‡ The ratio detector performs the functions of both a dynamic limiter and a frequency demodulator and is thus used in many entertainment-type FM receivers. The limiting operation is accomplished at the expense of balancing, and thus the signal component of

---

† To show that $R_{\text{in}} = R_1 \| (R/4)$ let $i_i = I_1 \cos \omega_0 t$ with the result that $v(t) = A \cos \omega_0 t$. If envelope detector loading is neglected, then $2v_b(t) = 2A \cos(\omega_0 t - \pi/2)$. Now the input power to the tuned transformer is given by $P_{\text{in}} = A^2/2R_{\text{in}}$. This power is dissipated in R_1 and R; thus $P_{\text{in}} = P_{R1} + P_R$. However, $P_R = 4A^2/2R$ and $P_{R1} = A^2/2R_1$; therefore,

$$\frac{1}{R_{\text{in}}} = \frac{1}{R_1} + \frac{4}{R} \qquad \text{or} \qquad R_{\text{in}} = R_1 \| \frac{R}{4}.$$

If envelope detector loading is not neglected, then

$$R_{\text{in}} = R_1 \| \frac{R}{4} \| \frac{R_0}{4}.$$

‡ See S. W. Seeley and J. Avins, "The Ratio Detector," *RCA Review*, **8**, p. 289 (1947). See also B. D. Loughlin, "The Theory of Amplitude Modulation Rejection in the Ratio Detector," *Proc. IRE*, **40**, pp. 289–296 (March 1952).

(t) has half the amplitude of the ouput of a corresponding Foster-Seeley demodu-
tor in addition to a dc component.

To get some idea of the operation of the ratio detector, consider the circuit of ig. 12.5-13 with C_2 removed. For this case, if we again assume that the loading of e envelope detectors can be modeled by placing R_0 in parallel with R, that operation such that the tuned transformer transfer function is given by Eq. (12.5–21), and that $_t/2 = 1$, then the upper and lower branches of the ratio detector function as ngle-ended time-delay demodulators, with the result that (cf. Eq. 12.5–16)

$$v_o'(t) = A\sqrt{2}[1 - (\Delta\omega t_0/2) f(t - t_0/2)],$$

$$v_o(t) = A\sqrt{2}[1 + (\Delta\omega t_0/2) f(t - t_0/2)], \qquad (12.5\text{–}24)$$

nd

$$v_2(t) = v_o'(t) + v_o(t) = 2A\sqrt{2}.$$

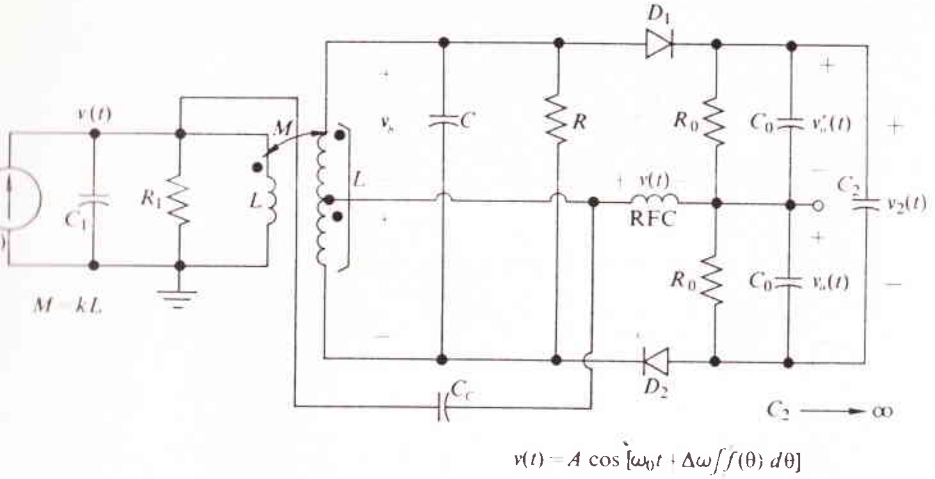


Fig. 12.5-13 Ratio detector.

Now if C_2 is returned to the circuit, it charges up to a point where the voltage oss it has the value

$$v_2(t) = 2A\sqrt{2}, \qquad (12.5\text{–}25)$$

which level no additional current flows through C_2 and the circuit functions as i were not there; that is, $v_o(t)$ is still given by Eq. (12.5–24). However, in a fashio ilar to that discussed in Section 12.1, C_2 placed across the secondary of the tune nsformer in series with diodes D_1 and D_2 performs the function of a dynami iter and keeps the amplitude of $v(t)$ and $v_b(t)$ constant at A even if $i_i(t)$ contains a ll amount of amplitude modulation. However, the modulation index (m) cannot too large; otherwise, failure-to-follow distortion will result.

In the case of the ratio detector it should be noted that the amplitude A of $v(t)$ is determined by the average amplitude of $i_i(t)$ flowing into the tuned transformer with C_2 removed. Thus, once the average amplitude I_1 of $i_i(t)$ and the parameters of the tuned circuit are known, A and in turn $v_o(t)$ can be calculated in a straightforward fashion:

$$A = I_1 R_1 \| \frac{R}{4} \| \frac{R_0}{4}.$$

In many practical ratio detector circuits, small resistors (small in comparison with R_0) are added in series with the diodes. These resistors serve two purposes. First, they provide a way of correcting for slight imbalances in the diodes and the center tapped transformer; second, they somewhat reduce the tendency of the circuit toward diagonal clipping.

Time-Delay Demodulator with a Synchronous Detector

The time-delay differentiator may be followed by a synchronous detector to yield an FM demodulator. Such an arrangement is shown in Fig. 12.5–14. For this demodulator the top branch is superfluous, since $v(t)$ multiplied by $-V_1 \sin \omega_0 \tau(t)$ produces

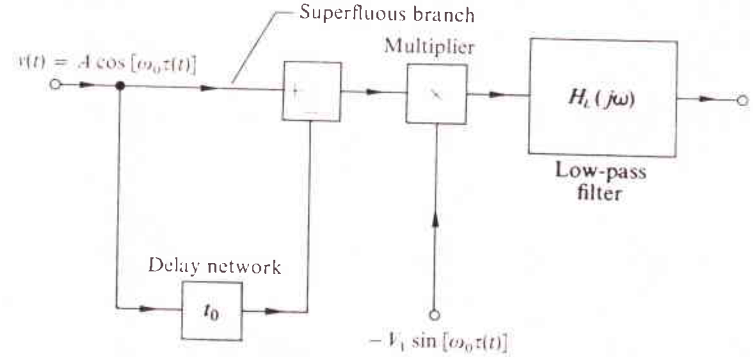


Fig. 12.5–14 Frequency demodulator comprising a time-delay differentiator followed by a synchronous detector.

a term centered about $2\omega_0$ which is not passed by the low-pass filter. In addition, if a phase shift of $-\pi/2$ at ω_0 is incorporated into the delay network and $v(t)$ is used as the reference signal, the output $v_0(t)$ remains unchanged; thus the simplified circuit of Fig. 12.5–15 produces the same output as the circuit of Fig. 12.5–14.

It is apparent for the circuit of Fig. 12.5–15 that the low-frequency component at the filter input is given by

$$-\frac{KA^2}{2} \sin 2k(t),$$

which simplifies to

$$-\frac{KA^2}{2}\Delta\omega t_0 f\left(t-\frac{t_0}{2}\right) \qquad (12.5\text{–}26)$$

provided that $t_0 < 0.2/\Delta\omega$ and $t_0 < 2/\omega_m$. If the low-pass filter extracts this low-frequency term (which is possible only if D and β determine a point below the curve of Fig. 12.2–6), then

$$v_o(t) = -K_M A\,\Delta\omega t_0 f\left(t-\frac{t_0}{2}\right), \qquad (12.5\text{–}27)$$

where $K_M = KA/2$. Equation (12.5–27) could have been obtained by letting $\theta_0 = 0$ in Eq. (12.5–12) and placing the resultant signal through a synchronous detector.

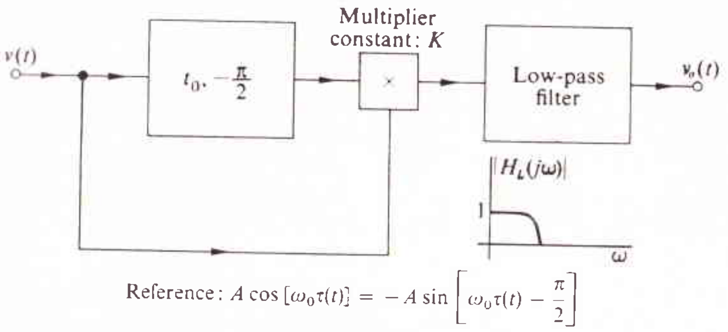


Fig. 12.5–15 Simplified time-delay frequency demodulator.

To arrive at Eq. (12.5–27) we require $t_0 < 0.2/\Delta\omega$ instead of $t_0 < 0.4/\Delta\omega$ as iven by Eq. (12.5–13). This is a direct result of employing $-V_1 \sin\omega_0\tau(t)$ as the eference signal for the synchronous detector in place of a more complex signal roportional to the high-frequency term of Eq. (12.5–3). Fortunately, the simplication of the reference signal has only this small effect.

Note that $v_o(t)$ given by Eq. (12.5–27) contains no dc component dependent on A. hus the frequency demodulator of Fig. 12.5–15 has the same insensitivity to amplide variations as any of the balanced frequency demodulators previously discussed. addition, it is easier to implement physically.

Figure 12.5–16 illustrates one possible physical implementation of the block iagram of Fig. 12.5–15. In this circuit, as in the circuit of Fig. 12.5–9, the transformer ith a tuned secondary is employed to achieve a time delay of $t_0 = 1/\alpha$ as well as a ase shift of $-\pi/2$ at the carrier frequency. Again the bandwidth of the input M signal

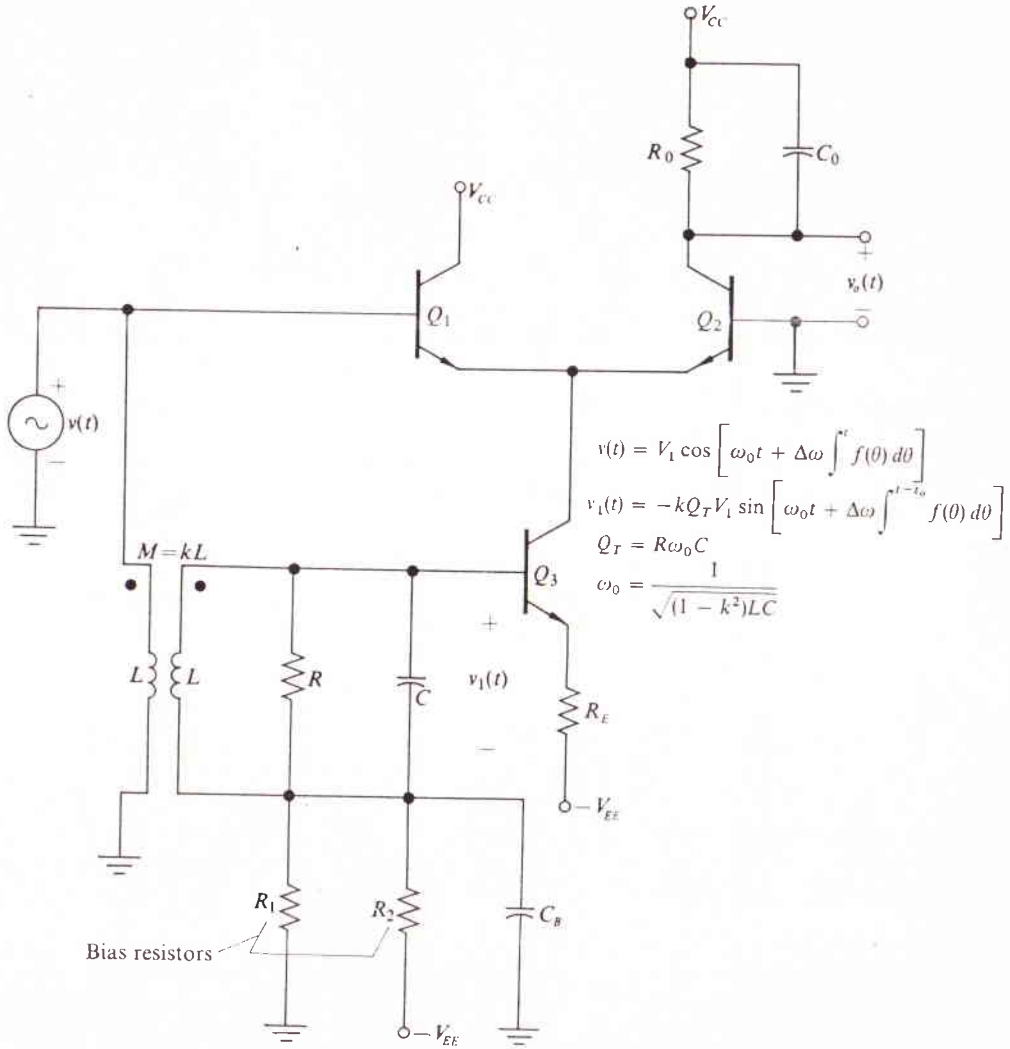
$$v(t) = V_1\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right]$$

must be less than 0.4α (cf. Eq. 12.5–23) and

$$t_0 = \frac{1}{\alpha} < \frac{0.2}{\Delta\omega}.$$

The differential pair is used as a multiplier, with the result that the low-frequency component of the collector current of Q_2 is given by

$$\frac{-\alpha_T^2 k Q_T V_1 a_1(x)}{2R_E\alpha}\Delta\omega f\left(t-\frac{1}{2\alpha}\right) + \alpha_T^2 I_{ko}, \qquad (12.5\text{–}28)$$



V_{CC}
R_0
C_0
V_{CC}
$v_o(t)$
Q_1
Q_2
+
−
v(t)
M = kL
L
L
R
C
Q_3
+
$v_1(t)$
−
R_E
$-V_{EE}$
R_1
R_2
C_B
Bias resistors
$-V_{EE}$
$v(t) = V_1\cos\left[\omega_0 t + \Delta\omega\int^t f(\theta)\,d\theta\right]$
$v_1(t) = -kQ_T V_1\sin\left[\omega_0 t + \Delta\omega\int^{t-t_0} f(\theta)\,d\theta\right]$
$Q_T = R\omega_0 C$
$\omega_0 = \frac{1}{\sqrt{(1-k^2)LC}}$

Fig. 12.5–16 Physical implementation of the block diagram of Fig. 12.5–15.

here α_T is the transistor alpha, $a_1(x)$ is the function of x given by Fig. 4.6–4, $x = V_1/kT$, and I_{k0} is the quiescent emitter current of Q_3 with $v(t) = v_1(t) = 0$. If the R_0C_0 low-pass filter extracts the low-frequency component of the collector current of Q_2 while eliminating the components centered about ω_0 and its harmonics, then $v_o(t)$ is given by

$$v_o(t) = V_{CC} - \alpha_T^2 R_0 I_{k0} + \frac{\alpha_T^2 R_0 k Q_T V_1 a_1(x)\,\Delta\omega f(t - 1/2\alpha)}{2R_E\alpha}. \qquad (12.5\text{–}29)$$

Of course, Eq. (12.5–29) is valid only if the bandwidth of the input FM signal is less han 0.4α and $t_0 = 1/\alpha \leq 0.2/\Delta\omega$. The stricter of the two conditions determines the maximum value of t_0 and, in turn, the largest amplitude for $v_o(t)$.

Bilotti[†] has reported an integrated 10.7 MHz differential-pair detector based on this approach. He also obtained an experimental output characteristic for his circuit.

2.6 PULSE-COUNT FREQUENCY DEMODULATOR

A pulse-count frequency demodulator is a time-delay frequency demodulator for which the time-delay network is a physical delay line, the amplitude demodulator is an average envelope detector, and the input signal is a hard-limited FM square wave $v_{sq}(t)$ as described by Eq. (12.2–1b). The basic block diagram of a pulse-count demodulator is shown in Fig. 12.6–1. Figure 12.6–2 illustrates the waveforms of the signals appearing at the various points of the block diagram of Fig. 12.6–1. From Fig. 12.6–2 we observe that a pulse of duration t_0 is generated in $v_a(t)$ at each positive zero crossing of $v_{sq}(t)$ and then averaged (or counted) by the low-pass filter to obtain $v_o(t)$.

Intuitively the operation of the pulse-count demodulator makes good sense. When the instantaneous frequency is high, the pulses of $v_a(t)$ are closely spaced and the average value (the low-frequency component) of the pulses at the low-pass filter output is high. Similarly, as the instantaneous frequency decreases, the pulses of $v_a(t)$ become more widely spaced and the low-pass filter output decreases. Thus the low-pass filter provides an output which is proportional to instantaneous frequency.

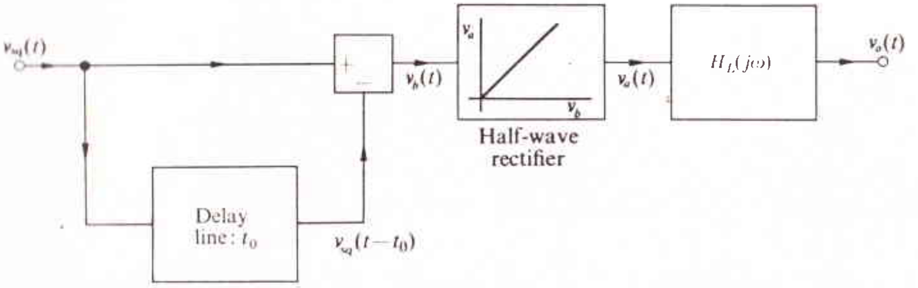


Fig. 12.6–1 Block diagram of pulse-count frequency demodulator.

A. Bilotti, "FM Detection Using a Product Detector," *Proc. IEEE*, **56**, pp. 755–757 (April 196

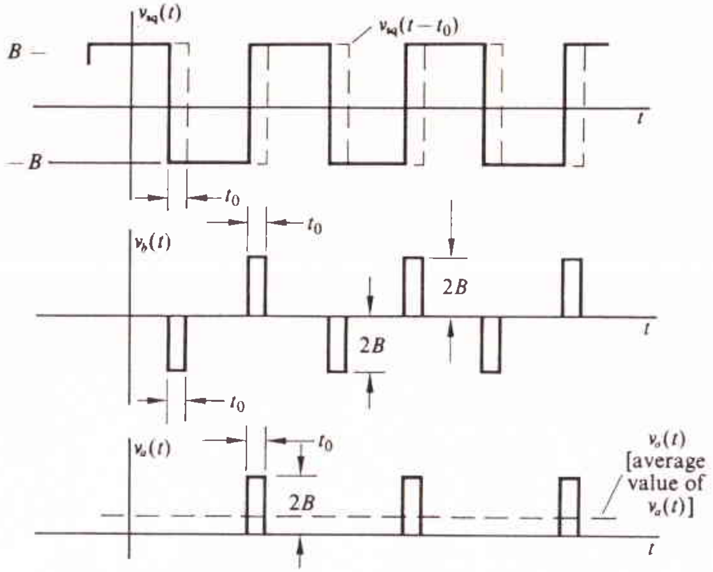


Fig. 12.6–2 Waveforms for the block diagram of Fig. 12.6–1.

For slowly varying $\omega_i(t)$ or $f(t)$, the constant of proportionality between $v_o(t)$ and $\omega_i(t)$ can readily be determined by letting $\omega_i(t) = \omega$ (a constant). For this case, $v_a(t)$ is a periodic train of pulses with an amplitude $2B$, with a duration t_0, with a period $2\pi/\omega$, and thus with an average value $Bt_0\omega/\pi$. Therefore,

$$v_o(t) = \frac{Bt_0\omega}{\pi} H_L(0), \qquad (12.6\text{–}1)$$

where $H_L(0)$ is the dc transfer function of the low-pass filter. A plot of $v_o(t)$ vs. ω is shown in Fig. 12.6–3. However, the voltage $v_o(t)$ is proportional to ω only for $T \geq 2t_0$. As T is decreased beyond this level, the pulses of $v_b(t)$ and $v_a(t)$ begin to decrease in duration with decreasing T, with the result that $v_o(t)$ decreases with increasing ω (for $\omega > \pi/t_0$) as shown in Fig. 12.6–3. [The reader should convince himself of this

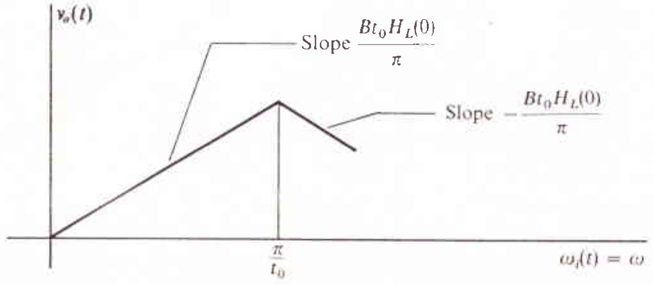


Fig. 12.6–3 Transfer characteristic for the pulse-count demodulator.

ct by subtracting $v_{sq}(t - t_0)$ from $v_{sq}(t)$ for the case where $T/2 < t_0$.] Because of e discontinuity in the $v_o(t)$-vs.-$\omega_i(t)$ curve, t_0 must be chosen sufficiently small so at

$$\frac{\pi}{t_0} \geq \omega_i(t)_{\max} = \omega_0 + \Delta\omega, \tag{12.6–2}$$

equivalently,

$$t_0 \leq \frac{\pi}{\omega_0 + \Delta\omega} \tag{12.6–3}$$

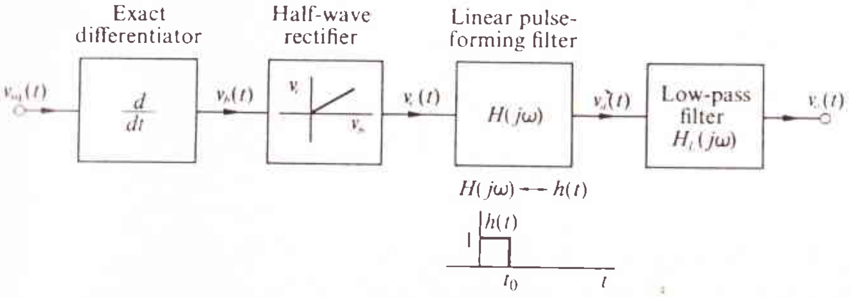
r any given value of $\omega_0 + \Delta\omega$.

The transfer characteristic of Fig. 12.6–3 was obtained on the assumption that (t) was a constant. We must now determine how rapidly $f(t)$ may vary relative to t_0 without altering the static transfer characteristic (or equivalently, how small t_0 ust be relative to a given ω_m in order not to alter the static transfer characteristic). o accomplish this we observe that $v_a(t)$ and thus $v_o(t)$ for the pulse-count demodutor and the circuit shown in Fig. 12.6–4 are identical, provided that the linear pulse-rming filter in Fig. 12.6–4 has an impulse response $h(t)$ given by

$$h(t) = \begin{cases} 0, & t < 0, \\ 1, & 0 \leq t \leq t_0, \\ 0, & t > t_0, \end{cases} \tag{12.6–4}$$

equivalently a transfer function $H(j\omega) = \mathscr{F}[h(t)]$ of the form

$$H(j\omega) = t_0 e^{-j\omega t_0/2} \frac{\sin(\omega t_0/2)}{\omega t_0/2}. \tag{12.6–5}$$



. 12.6–4 Block diagram of a circuit whose output is identical with that of the circuit of 12.6–1.

e equivalence is apparent when we note that the exact differentiator produces an ernating train of positive and negative impulses of strength $2B$ at the respective itive and negative zero crossings of $v_{sq}(t)$. The positive impulses, which get past half-wave rectifier of Fig. 12.6–4, are converted into rectangular pulses of height and duration t_0 by the linear pulse-forming filter.

The pulse-forming filter has a magnitude and phase response identical with that shown in Fig. 12.5–4 except that $|H(0)|$ equals t_0 instead of $\Delta\omega t_0/2$. However, this filter is in cascade with the output low-pass filter which cuts off immediately beyond ω_m [the maximum frequency component of $f(t)$]; hence, if

$$t_0 < \frac{2}{\omega_m} \tag{12.6–6}$$

so that $|H(j\omega)|$ remains essentially constant at t_0 over the passband of the output low-pass filter, the linear pulse-forming filter may be replaced by an attenuator of value t_0 without affecting $v_o(t)$. This substitution reduces the circuit of Fig. 12.6–4 to an exact differentiator followed by an average detector and an attenuator of t_0; hence with the aid of Eq. (12.2–17) ($K_D = 1$) we can write $v_o(t)$ in the form

$$v_o(t) = \frac{Bt_0}{\pi}[\omega_0 + \Delta\omega f(t)] * h_L(t) \tag{12.6–7}$$

or

$$v_o(t) = \frac{Bt_0 H_L(0)}{\pi}\omega_i(t),$$

provided that the output filter passes $v_o(t)$ without distortion.

Clearly, then, if t_0 is chosen sufficiently small so that both Eqs. (12.6–6) and (12.6–3) are satisfied, the constant of proportionality relating $v_o(t)$ and $\omega_i(t)$ is $Bt_0H_L(0)/\pi$ both dynamically and statically. Thus, once ω_0, $\Delta\omega$, and ω_m are specified for the input FM signal and β and D are found to be below the curve of Fig. 12.2–7 [so that the low-frequency component of $v_o(t)$ can be extracted by the output low-pass filter], then t_0 may be chosen sufficiently small so that $v_o(t)$ is linearly related to $\omega_i(t)$. For the usual case where $\omega_m \ll \omega_0$, Eq. (12.6–3) is more restrictive than Eq. (12.6–6) and thus we need only concern ourselves with choosing t_0 sufficiently small so that the static demodulator transfer characteristic remains linear. Of course, t_0 should be chosen no smaller than necessary, since the amplitude of $v_o(t)$ is directly proportional to t_0.

A practical system which performs the same operation as the time-delay differentiator followed by the average envelope detector is the pulse-count frequency demodulator shown in Fig. 12.6–5. For this demodulator the monostable multivibrator is triggered on the positively sloping edge of $v_{sq}(t)$ and produces a pulse of amplitude V_2 and duration t_0. The low-pass filter then extracts the low-frequency component of the pulse train.

The static transfer characteristic of $v_o(t)$ vs. $\omega_i(t)$ (which is valid provided that $t_0 < 2/\omega_m$) is shown in Fig. 12.6–6. For this case no reduction in pulse duration occurs until $T = 2\pi/\omega$ decreases to the point where the monostable multivibrator cannot recover between the end of the pulse of duration t_0 and the beginning of the next pulse. If the time for complete recovery of the monostable multivibrator is t_R, then the minimum permissible size of T is $T = t_0 + t_R$. Consequently, in order to obtain

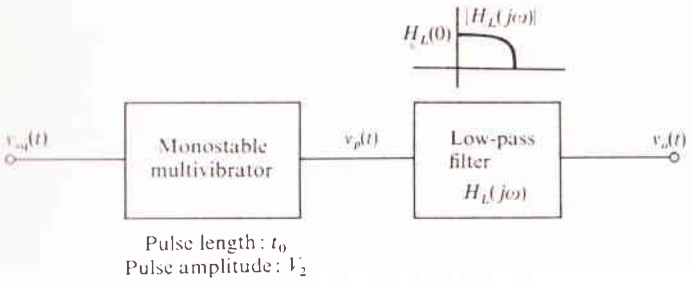


Fig. 12.6–5 Pulse-count demodulator.

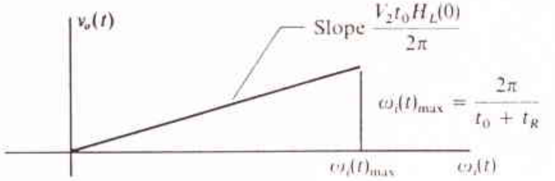


Fig. 12.6–6 Plot of the transfer characteristic for the pulse-count frequency demodulator shown in Fig. 12.6–5.

the desired linear relationship between $v_o(t)$ and $\omega_i(t)$,

$$v_o(t) = \frac{V_2 t_0 H_L(0)}{2\pi}\,\omega_i(t). \tag{12.6–8}$$

t_0 must be chosen sufficiently small so that

$$t_0 + t_R \leq \frac{2\pi}{\omega_i(t)_{\max}} = \frac{2\pi}{\omega_0 + \Delta\omega} \tag{12.6–9}$$

and

$$t_0 < \frac{2}{\omega_m}. \tag{12.6–10}$$

The validity of Eq. (12.6–8) depends, of course, also on the ability of the low-pass output filter to extract the low-frequency component of the pulse train in an undistorted fashion.

Commercial discriminators of this type are available as separate units. Center frequencies from 1 Hz to 10 MHz are available. Above 10 MHz the pulse duration must be less than 100 nsec and the recovery time must be of the order of 10 nsec; hence practical circuit considerations impose limitations. When operated within the previously described limitation, such circuits are capable of achieving significantly better than 0.1 % linearities with frequency deviations approaching the carrier frequency.

Commercial divide-by-two and divide-by-ten networks are now available which operate with input frequencies as high as 1000 MHz. With such counting networks the useful frequency range of the pulse-count discriminator can be extended considerably in frequency.

For cases where D lies reasonably close to the lower curve of Fig. 12.2–7 the modulation cannot easily be separated from the carrier ripple without a reasonably complex low-pass filter. To remedy this situation somewhat, the monostable multivibrator of Fig. 12.6–5 may be triggered on *every* zero crossing of $v_{sq}(t)$ rather than only on positive zero crossings of $v_{sq}(t)$. With such triggering the closest carrier frequency components to the desired modulation at the input to the low-pass filter result from the FM spectrum centered at $2\omega_0$ rather than at ω_0.

Therefore, the upper curve of Fig. 12.2–7 provides a bound for D. If the monostable multivibrator is triggered on every zero crossing, the output voltage of the pulse-count demodulator is given by

$$v_o(t) = \frac{V_2 t_0 H_L(0)}{\pi}\,\omega_i(t), \tag{12.6–11}$$

from which we observe that the output level has apparently been increased by a factor of 2. This increase cannot normally be realized, since for this case

$$t_0 + t_R \leq \frac{1}{\omega_0 + \Delta\omega}, \tag{12.6–12}$$

which constrains the maximum permissible value of t_0 to be less than one-half the previous value.

12.7 MORE EXOTIC FM DETECTORS—THE PHASE-LOCKED LOOP, THE FREQUENCY-LOCKED LOOP, AND THE FREQUENCY DEMODULATOR WITH FEEDBACK

As the final section of this chapter and of the book we present the block diagrams for and brief résumés of the operation of three "threshold-extending" FM demodulators. All of these circuits are more complex than the demodulators discussed previously; however, since all the individual elements within each block diagram have been considered in this book, the interested reader should have an adequate basis for understanding their operation.

Because this section makes no pretense of offering a complete explanation of these useful but not yet widely used circuits, we have included a fairly substantial list of references at the end of the chapter which will supply further information. Even more recent information will probably be available from an inspection of recent indexes to papers in the *IEEE Transactions on Communication Technology or Aerospace and Electronic Systems.*

Figures 12.7–1, 12.7–2, and 12.7–3 show all three circuits in block diagram form. All three circuits contain feedback loops, all contain multipliers within the loop, and all three take their outputs after a low-pass filter that is within the loop. If preceded by a limiter, the $a(t)$ amplitude terms shown would become constants;

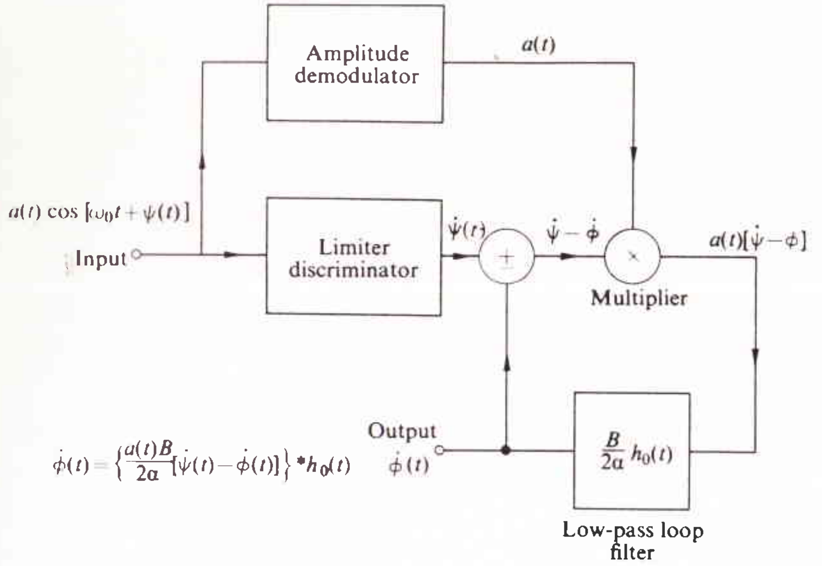

Amplitude demodulator
$a(t)$
$a(t)\cos[\omega_0 t + \psi(t)]$
Input
Limiter discriminator
$\dot{\psi}(t)$
$\dot{\psi} - \dot{\phi}$
$a(t)[\dot{\psi} - \dot{\phi}]$
Multiplier
$\dot{\phi}(t) = \left\{\frac{a(t)B}{2a}[\dot{\psi}(t) - \dot{\phi}(t)]\right\} * h_0(t)$
Output
$\dot{\phi}(t)$
$\frac{B}{2a} h_0(t)$
Low-pass loop filter

Fig. 12.7–1 Block diagram of frequency-locked loop.

nd though the circuits would detect FM signals uncorrupted by noise, they would se their ability to remove or reduce the effects of noise-induced "spikes" and hence eir ability to extend the FM noise threshold. It is just this ability to extend the reshold that justifies the added complexity of these circuits in various weak-signal M situations—satellite links, for example.

The frequency-locked loop combines an envelope detector and a feedback loop a circuit to process the output from a conventional limiter-discriminator circuit. long as the input signal is large, the envelope detector output keeps the loop in high and hence the loop bandwidth wide and the discriminator output reaches

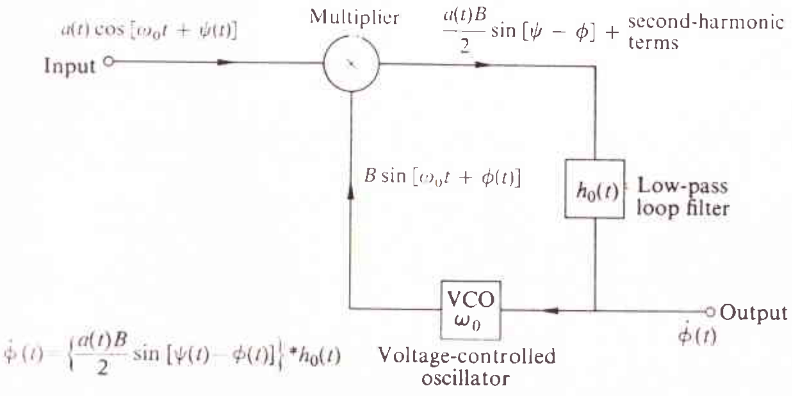

$a(t)\cos[\omega_0 t + \psi(t)]$
Input
Multiplier
$\frac{a(t)B}{2}\sin[\psi - \phi]$ + second-harmonic terms
$B\sin[\omega_0 t + \phi(t)]$
$h_0(t)$
Low-pass loop filter
VCO
ω_0
Output
$\phi(t)$
$\phi(t) = \left\{\frac{a(t)B}{2}\sin[\psi(t) - \phi(t)]\right\} * h_0(t)$
Voltage-controlled oscillator

Fig. 12.7–2 Block diagram of phase-locked loop.

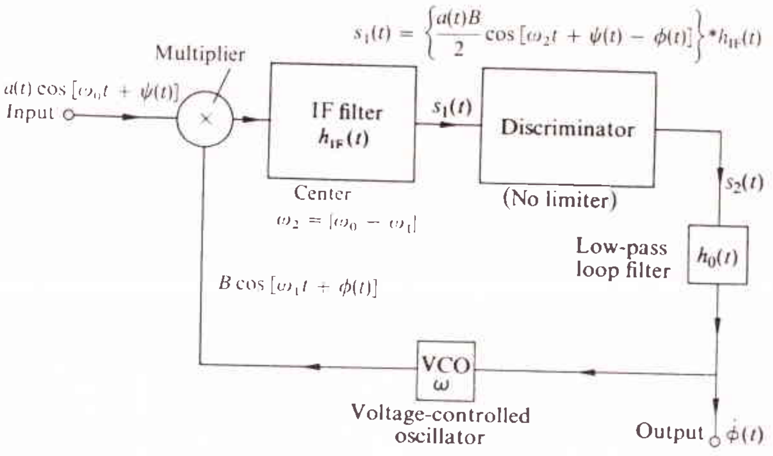

$s_1(t) = \left\{\frac{a(t)B}{2}\cos[\omega_2 t + \psi(t) - \phi(t)]\right\} * h_{IF}(t)$
Multiplier
$a(t)\cos[\omega_0 t + \psi(t)]$
Input
IF filter
$h_{IF}(t)$
$s_1(t)$
Discriminator
(No limiter)
$s_2(t)$
Center
$\omega_2 = |\omega_0 - \omega_1|$
$B\cos[\omega_1 t + \phi(t)]$
Low-pass loop filter
$h_0(t)$
VCO
ω
Voltage-controlled oscillator
Output
$\dot{\phi}(t)$

Fig. 12.7–3 Block diagram of frequency demodulator with feedback.

the circuit output unchanged. When the input amplitude drops briefly to a small amplitude (it can be shown that sudden drops to a small amplitude are often caused by the same noise mechanism that introduces "spikes" or impulselike disturbances in the FM signal), the loop gain drops and the circuit output is uncoupled from the discriminator output until the disturbance passes. Some of the references listed at the end of the chapter contain experimental results as well as much more complete discussions of the circuit operation.

As was pointed out previously, one approach to the phase-locked loop is to consider the voltage-controlled oscillator as an FM generator in the return path of a feedback loop. If $a(t)$ is large, then the loop gain is high and the forward path must try to perform the inverse of FM generation or it must approximate an FM detector. When $a(t)$ falls—again during a noise disturbance—the loop tends to open and the output tends to be unaffected by any noise spikes present in the input FM signal. As pointed out below, the phase-locked loop suffers from problems of initial locking on a signal, from VCO drift, and, if poorly designed, from unwanted "loss-of-lock" conditions.

The frequency demodulator with feedback includes both a voltage-controlled oscillator and a conventional discriminator within a feedback loop. Also included within the same loop are a low-pass filter and a single-pole IF filter. Hess† has shown that as the IF filter bandwidth goes to zero the equations for this circuit reduce to those for the phase-locked loop, whereas when the IF filter becomes wide and the low-pass filter becomes narrow the equations approach those for the frequency-locked loop.

Again an intuitive argument for the circuit operation is that, when the amplitude is high, the loop gain is high and the feedback widens the circuit bandwidth and esents the discriminator output unchanged. When $a(t)$ falls, then the loop gain nd circuit bandwidth fall rapidly and the output is uncoupled from the unwanted istortions in the input sine wave.

---

† D. T. Hess, "Equivalence of FM Threshold Extension Receivers," *IEEE Trans. on Communication Technology*, **Com-16**, No. 5, pp. 946–748 (Oct. 1968).

hase-Locked Loop Operation

To see mathematically how the simple phase-locked loop initially becomes ocked" to the incoming signal, consider the simplified case where an unmodulated arrier at $\omega_0 + \varepsilon$ is tuned slowly toward the free-running VCO frequency ω_0. Then in ig. 12.7–2 $\psi(t)$ is εt and at $t = 0$ one may assume the VCO output to be cosinusoidal. ne also assumes that the low-pass loop filter rejects additive frequency terms and asses difference terms with an attenuation independent of frequency. Under these onditions,

$$\dot{\phi}(t) = \omega_x \cos[\varepsilon t - \phi(t)], \qquad (12.7\text{–}1)$$

here ω_x contains $a(t)$ (assumed constant for the moment), B, the filter constant, e $\frac{1}{2}$ term from the multiplication, and the multiplier constant; ω_x is known as the oop bandwidth" or the loop "hold-in range."

Now Eq. (12.7–1) is a first-order nonlinear differential equation which can be lved, on the assumption that $\phi = 0$ at $t = 0$, to yield

$$\dot{\phi}(t) = \frac{\varepsilon^2 - \omega_x^2}{\varepsilon - \omega_x \cosh\sqrt{\omega_x^2 - \varepsilon^2 t}} - \varepsilon, \qquad |\varepsilon| < \omega_x, \qquad (12.7\text{–}2)$$

$$\dot{\phi}(t) = \frac{\varepsilon^2 - \omega_x^2}{\varepsilon - \omega_x \cos\sqrt{\varepsilon^2 - \omega_x^2 t}} - \varepsilon, \qquad |\varepsilon| > \omega_x. \qquad (12.7\text{–}3)$$

hese two equations are plotted in Figs. 12.7–4 and 12:7–5 for positive values of ω_x.

When $\omega_x > |\varepsilon|$, then as time passes $\dot{\phi}$ approaches ε and the loop is "locked." ε approaches zero, then $\dot{\phi}$ also approaches zero. From Fig. 12.7–4 it is apparent at it takes about five "time constants" equal to $1/\sqrt{\omega_x^2 - \varepsilon^2}$ for this locking to ccur.

Now, since in the steady state $\dot{\phi}(t) = \varepsilon$, if ε is varied slowly in comparison with $\sqrt{\omega_x^2 - \varepsilon^2}$, then the output of the phase-locked loop produces an output directly roportional to the frequency difference between the input and the VCO frequency. onsequently, if

$$\psi(t) = \Delta\omega \int^t f(\theta)\, d\theta,$$

the maximum frequency component of $f(t)$ is ω_m, and if

$$\omega_m \ll \sqrt{\omega_x^2 - \Delta\omega^2}, \qquad (12.7\text{–}4)$$

en the output of the phase-locked loop is directly proportional to $f(t)$ and frequency demodulation is accomplished.

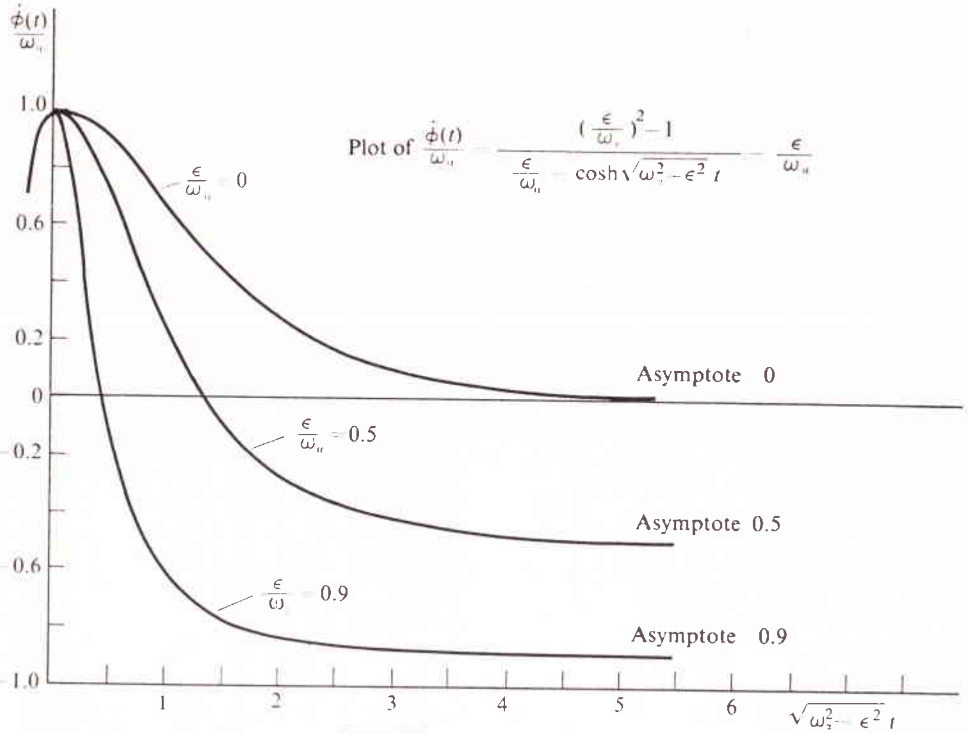


Fig. 12.7–4 Plot of $\dot{\phi}(t)/\omega_x$ vs. $\sqrt{\omega_x^2 - \varepsilon^2 t}$ for $|\varepsilon/\omega_x| < 1$. The "lock-in" range for a first-order phase-locked loop.

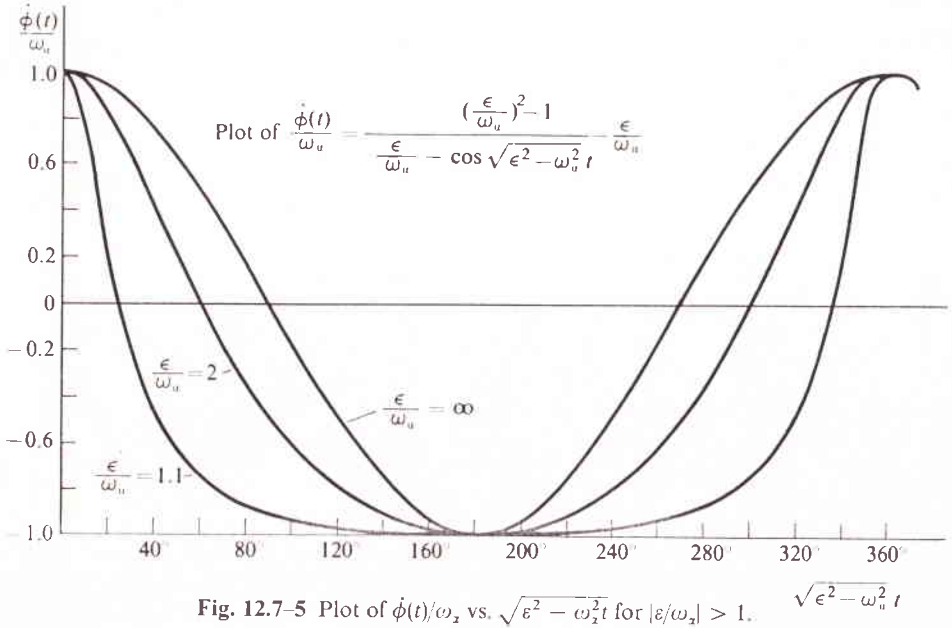


Fig. 12.7–5 Plot of $\dot{\phi}(t)/\omega_x$ vs. $\sqrt{\varepsilon^2 - \omega_x^2 t}$ for $|\varepsilon/\omega_x| > 1$.

The inequality of Eq. (12.7–4) implies that $\omega_x > \Delta\omega$ and that $\omega_x \gg \omega_m$. In addition, it can be shown that, in order to avoid distortion in the detected signal, D and β must fall below the curve of Fig. 12.2–6; hence the phase-locked loop detector suffers from the same type of limitations concerning carrier frequency ω_0, modulation frequency ω_m, and frequency deviation $\Delta\omega$ as do the other demodulators discussed in this chapter.

When $|\varepsilon| > \omega_x$, then $\dot{\phi}(t)$ goes on oscillating with time at a beat frequency dependent on $\sqrt{\varepsilon^2 - \omega_x^2}$. In this case locking never occurs unless one either decreases ε or increases the "loop bandwidth" ω_x.

Real phase-locked loops have more complex filters and have "pull-in" and "hold-in" ranges that are different. (The "hold-in" range exceeds the "pull-in" or "lock-in" range.) However, in all cases these ranges are a function of the input signal amplitude. Thus, if the amplitude suddenly dips, the circuit may "lose lock"; if this happens, the output will be unaffected by the input frequency disturbances that accompany such suddenly dipping amplitudes.

REFERENCES

L. Calandrino and G. Immovilli, "Coincidence of Pulses in Amplitude and Frequency Deviations," *Alta Frequenza*, **36**, English Issue No. 3 (Aug. 1967).

F. Cassara, "FM Demodulator with Feedback," M.S. Project Report, Polytechnic Institute of Brooklyn (June 1968).

L. H. Enloe, "Decreasing the Threshold in FM by Frequency Feedback," *Proc. IRE*, **50**, pp. 18–30 (Jan. 1962).

P. Frutiger, "Noise in FM Receivers with Negative Frequency Feedback," *Proc. IEEE*, **54**, pp. 1506–1520 (Nov. 1966).

D. T. Hess, "Cycle Slipping in First Order Phase Locked Loops," *IEEE Trans. on Communication Technology*, **Com-16**, No. 2, pp. 255–260 (April 1968).

D. T. Hess, "Equivalence of FM Threshold Extension Receivers," *IEEE Trans. on Communication Technology*, **Com-16**, No. 5, pp. 746-748 (Oct. 1968).

D. T. Hess and K. K. Clarke, "Optimization of the Frequency Locked Loop," *Proc. National Electronics Conference*, pp. 528–533 (Dec. 1969).

S. O. Rice, "Noise in FM Receivers," in *Time Series Analysis*, M. Rosenblatt, Ed., Wiley, New York (1963), Chapter 25.

M. Schwartz, W. R. Bennett, and S. Stein, *Communication Systems and Techniques*, McGraw-Hill, New York (1966), Chapter 3.

M. Unkauf, "Quantized Frequency Locked Loop FM Demodulator," Ph.D. Dissertation, Polytechnic Institute of Brooklyn (June 1969).

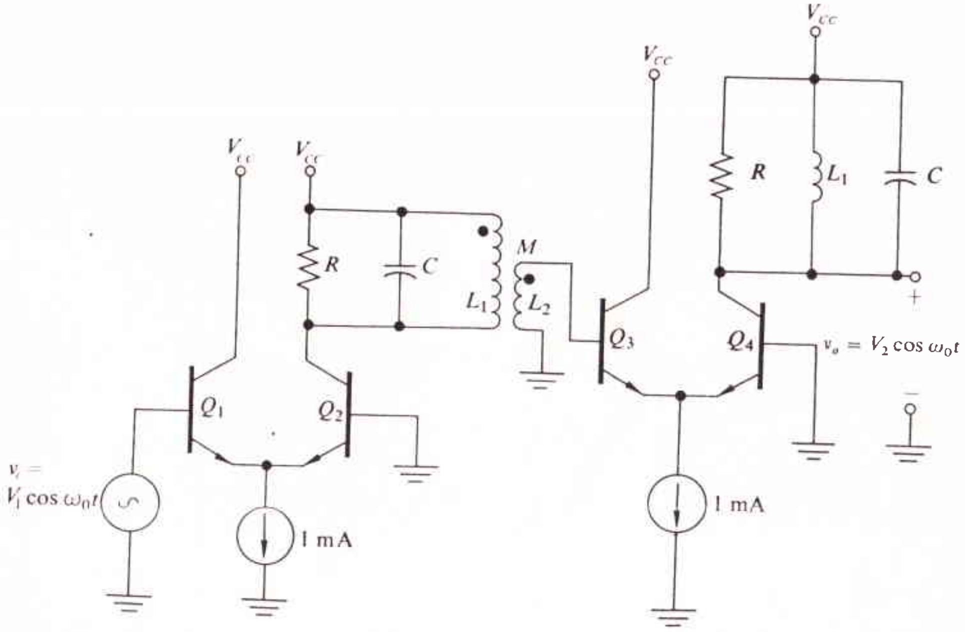
PROBLEMS

12.1 For the circuit shown in Fig. 12.P–1, plot the output voltage amplitude V_2 vs. the input voltage amplitude V_1 for $0 < V_1 < 50$ mV. At what value of V_1 is V_2 within 95% of its

asymptotic limiting value? If

$$v_1 = (1.6\text{ mV})(1 + 0.5\cos\omega_m t)\cos\omega_0 t, \qquad \omega_m \ll \omega_0,$$

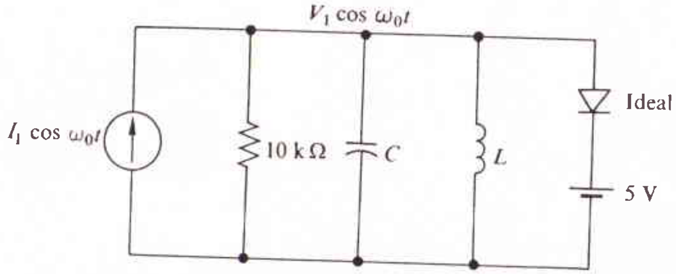
find the modulation index of $v_o(t)$.



$$\frac{M}{L_1} = 0.05, \quad \frac{1}{\sqrt{L_1 C}} = \omega_0, \quad \omega_0 RC = 20, \quad V_{CC} = 10\text{ V}, \quad \text{R} = 10\text{ k}\Omega$$

Figure 12.P–1

12.2 For the circuit shown in Fig. 12.P–2, plot the transfer characteristic of V_1 vs. I_1 for $0 \le I_1 < 5$ mA. Assume that the loaded Q remains sufficiently high to justify the assumption of a sinusoidal output voltage.



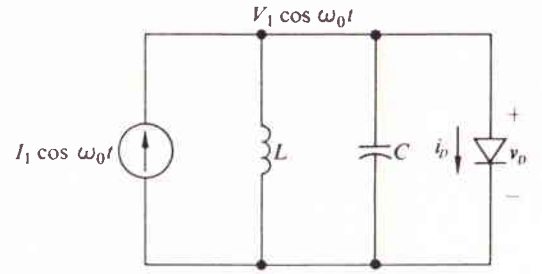
$\omega_0 = 1/\sqrt{LC}$

Figure 12.P–2

12.3 For the circuit shown in Fig. 12.P–3, show that if the loaded Q is high, the amplitude of the output voltage V_1 is related to the amplitude of the input current by the expression

$$\frac{I_1}{I_S} = 2I_1(x),$$

where $I_1(x)$ is the modified Bessel function of order one and $x = qV_1/kT$. Using the values for $I_1(x)$ and the asymptotic expression for $I_1(x)$ given in the Appendix at the back of the book, plot V_1 vs. I_1 for $0 \leq I_1 < 5$ mA. Does this circuit function as a limiter? For what range of I_1? What value of C ensures $Q_T = 20$ if $I_1 = 5$ mA and $\omega_0 = 10^8$ rad/sec?



$\omega_0 = 1/\sqrt{LC}$, $i_D = I_s e^{q v_D/kT}$, $I_S = 10^{-16}$ A

Figure 12.P–3

12.4 A high-Q single-tuned circuit follows an ideal nonlinear limiter (Fig. 12.1–1) to extract the fundamental component of the FM signal. Assuming that the filter functions in a quasi-static fashion (Eq. 11.3–2), determine the minimum filter bandwidth (in terms of $\Delta\omega$) which ensures no more than 1% amplitude modulation on the output signal ($m = 0.01$). With $\beta = 5$, determine the maximum deviation ratio which is possible if the tuned circuit must suppress the third harmonic component of the limited FM signal to less than 1% of the fundamental while keeping the induced AM below 1%.

12.5 Repeat Problem 12.4 for the case where the bandpass filter is a maximally flat double-tuned circuit.

12.6 For the bandpass limiter shown in Fig. 12.P–4, $v_1(t) = V_1 \cos \omega_0 t$, which results in $v_o(t) = V_2 \cos \omega_0 t$. Using the sine-wave tip analysis in Sections 4.2 and 4.3, plot the limiter characteristic of V_2 vs. V_1. If

$$v_1(t) = (5 \text{ V})[1 + 0.1f(t)] \cos \omega_0 t$$

and the modulation is passed by the bandpass filter, determine the AM modulation index of $v_o(t)$.

12.7 For the dynamic limiter shown in Fig. 12.1–5, $L = 10\ \mu$H, $C = 1000$ pF, $R = R_0 = 10$ kΩ, $C_0 = 10\ \mu$F, and

$$i_i(t) = (1 \text{ mA})(1 + 0.5 \cos 10^3 t) \cos\left(10^7 t + 10^4 \int^t \cos 10^3\theta\, d\theta\right).$$

Find $v(t)$. What is the AM modulation index of $v(t)$? What is the output AM modulation index if C_0 is reduced to 1 μF? Does failure-to-follow distortion occur?

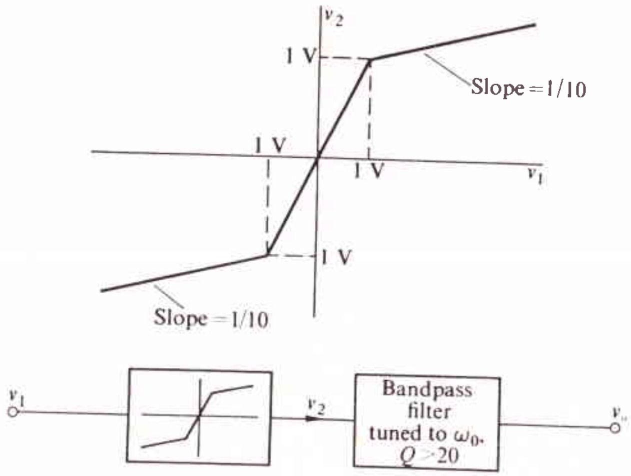


Figure 12.P–4

12.8 For the circuit shown in Fig. 12.P–5,

$$v_i(t) = (0.5 \text{ V})(1 + 0.5 \cos 2.5 \times 10^5) \cos 10^7 t.$$

Find $v_o(t)$. What is its AM modulation index? If

$$v_i(t) = (0.5 \text{ V}) \cos\left(10^7 t + \Delta\omega \int^t \cos 10^4\theta\, d\theta\right)$$

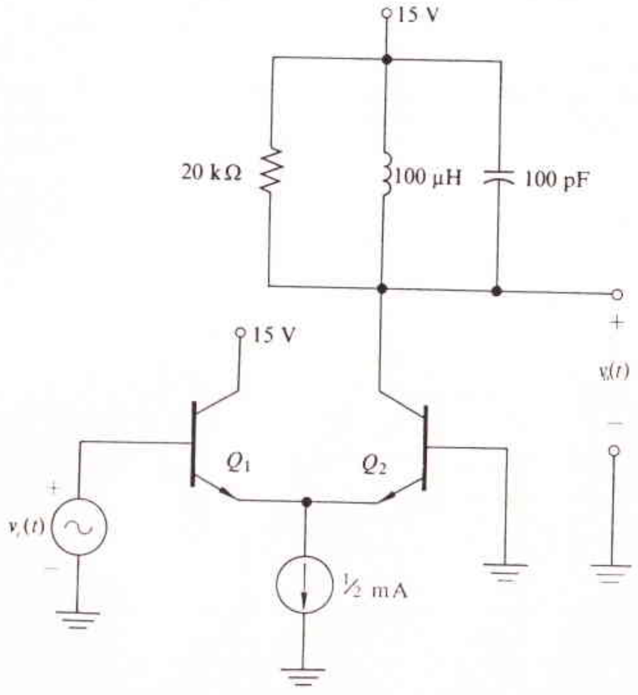


Figure 12.P–5

is applied to the circuit of Fig. 12.P–5, find the maximum value of $\Delta\omega$ which ensures less than 1 % amplitude modulation ($m = 0.01$) for $v_o(t)$.

12.9 The Clarke-Hess demodulator shown in Fig. 12.3–4 is to be used to demodulate an FM triangular wave. Show that in order to balance the demodulator for this case one must choose $RC = \pi/4\omega_0$. Show also that when the demodulator is balanced, the output is given by Eq. (12.3–7), where A is now the peak amplitude of the FM triangular wave.

2.10 For the Clarke-Hess demodulator shown in Fig. 12.3–4, $R = 1000\,\Omega$, $C = 100\,\text{pF}$, $R_0 = 15.7\,\text{k}\Omega$, $C_0 = 6300\,\text{pF}$, $V_{CC} = 10\,\text{V}$ and

$$v_i(t) = (10\,\text{V}) \cos \left(10^7 t + 10^6 \int^t \cos 10^4 \theta \, d\theta \right).$$

Find $v_o(t)$. How much carrier ripple appears at the output?

2.11 A 10 V peak-to-peak FM square wave with instantaneous frequency $\omega_i(t)$ is applied to the circuit of Fig. 12.3–4 with $C = 100\,\text{pF}$, $R_0 = 15.7\,\text{k}\Omega$, $C_0 = 6300\,\text{pF}$, and $V_{CC} = 10\,\text{V}$.

a) Find the value of R which balances the demodulator at $\omega_0 = 10^7$ rad/sec.

b) If $r_\Sigma = 100\,\Omega$ and $\omega_i(t) = \omega_0 = 10^7$ rad/sec, make an accurate sketch of the current flowing into the output R_0-C_0 filter. Determine the output carrier ripple.

c) Plot the v_o-vs.-$\omega_i(t)$ characteristic as $\omega_i(t)$ is varied between 0 and $2\omega_0$.

d) If $\omega_i(t) = (10^7 + 5 \times 10^4 \cos 10^3 t)$ rad/sec, determine $v_o(t)$.

12.12 For the circuit shown in Fig. 12.P–6, $v_i(t)$ is given by

$$v_i(t) = (2\,\text{V}) \cos \left[\omega_0 t + \Delta\omega \int^t f(\theta) \, d\theta \right],$$

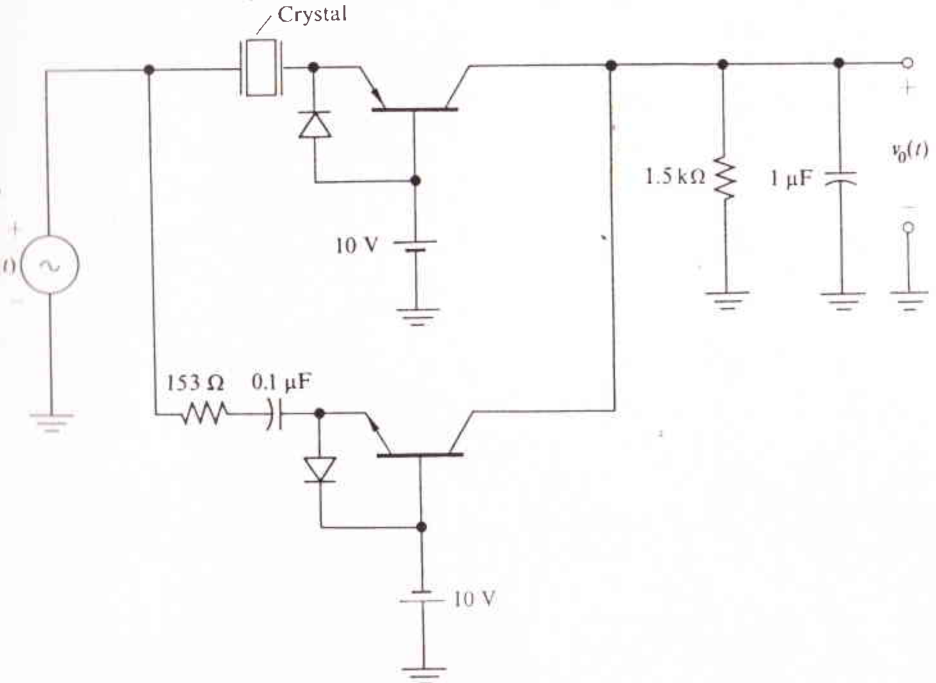


Figure 12.P–6

where $\omega_0 = (10^7 + 176)$ rad/sec. The crystal parameters are given by $C_0 = 4\,\text{pF}$, $C = 0.04\,\text{pF}$, $L = 250\,\text{mH}$, and $r = 125\,\Omega$ (cf. Section 6.7). Plot the static demodulator characteristic v_o vs. $\Delta\omega f(t)$ for $-500\,\text{rad/sec} \le \Delta\omega f(t) \le 500\,\text{rad/sec}$. (*Hint:* Since operation is in the vicinity of the crystal's series resonance, the effect of C_0 can be neglected.) What is the maximum permissible value of $\Delta\omega$ which ensures that the nonlinearity of the transfer characteristic does not exceed 1 %? With this value of $\Delta\omega$, with $\beta = 5$, and with $f(t) = \cos \omega_m t$, find $v_o(t)$. If R increases by 10 %, what is the change in the center frequency of the demodulator? Would this demodulator be useful in the system shown in Fig. 11.9–1? Why?

12.13 Verify the condition given in Eq. (12.4–14) and the expansion of Eq. (12.4–15).

12.14 For the balanced demodulator shown in Fig. 12.4–3, $R_1 = R_2 = 50\,\text{k}\Omega$ and $R_0 = 100\,\text{k}\Omega$. Choose optimum values for the remaining unspecified parameters if the circuit is to demodulate an input FM signal limited to 10 mA peak-to-peak having a carrier frequency $\omega_0 = 10^8$ rad/sec, a frequency deviation of $\Delta\omega = 10^6$ rad/sec, and a maximum modulation frequency of 10^5 rad/sec. Find $v_o(t)$ for this choice of parameters.

12.15 A ruby laser beam is modulated in frequency with a deviation of $\Delta f = 1$ GHz and a maximum modulation frequency of $f_m = 6$ MHz. Draw a diagram of an optimum frequency demodulator for such a beam. The demodulator should employ only beam splitters (prisms), polished mirrors, and a photodiode (or phototransistor). Indicate all effective path lengths, assuming that the carrier beam propagates at the speed of light through the demodulator medium.

If the received beam has an intensity of 30 foot-candles and if the photodiode has a linear sensitivity of 100 μA/foot-candle and delivers its output current to a 600 Ω impedance, determine the output voltage for the case where $f(t) = \cos(2\pi \times 6 \times 10^6 t)$.

12.16 In the bottom circuit of Fig. 12.5–8 the delay network is a piece of helically wound delay line with a delay of 20 nsec/cm and a length of 10 cm. If $v(t) = (5\,\text{V}) \cos \omega t$ and $K_M = 1$, plot the v_o-vs.-ω transfer characteristic for $0 < \omega < 5 \times 10^7$. Over what ranges of ω is this characteristic linear? Do these ranges of ω agree with the results of Section 12.5?

12.17 If the envelope detector in the demodulator shown in Fig. 12.P–7 is assumed not to load 75 Ω, show that $v_a(t)$ is given by

$$v_a(t) = \tfrac{1}{2}[v_i(t) - v_i(t - 2t_0)],$$

where $t_0 = 400$ nsec. If, in addition,

$$v_i(t) = (5\,\text{V}) \cos \left(\omega_0 t + \Delta\omega \int^t \cos \omega_m \theta \, d\theta \right),$$

where $\omega_0 = 3.93 \times 10^6$ rad/sec, $\Delta\omega = 10^6$ rad/sec, and $\beta = 10$, find $v_o(t)$. Is the demodulation nonlinearity less than 1 %?

12.18 A 10 V peak-to-peak square wave with frequency ω is applied to the demodulator of Fig. 12.P–7.

a) Sketch $v_a(t)$ for $\omega < \pi/2t_0$ and for $\omega > \pi/2t_0$.

b) Plot the v_o-vs.-ω characteristic for $0 < \omega < \pi/t_0$. Compare this characteristic with that shown in Fig. 12.6–3.

12.19 A 2 V peak-to-peak FM square wave $v_{sq}(t)$ triggers the monostable multivibrator shown in Fig. 12.P–8.

a) If the instantaneous frequency $\omega_i(t) = \omega$ (a constant), plot the v_o-vs.-ω characteristic for $0 < \omega < 10^6$ rad/sec. At what point does the characteristic cease to be linear?

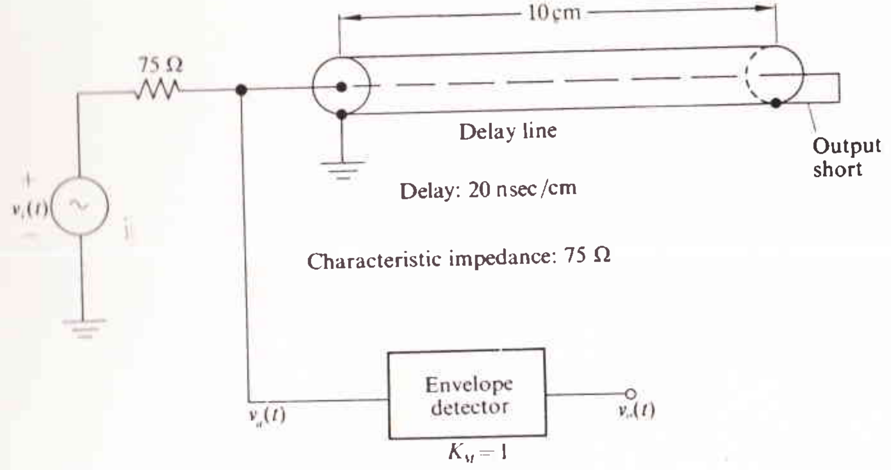


Figure 12.P–7

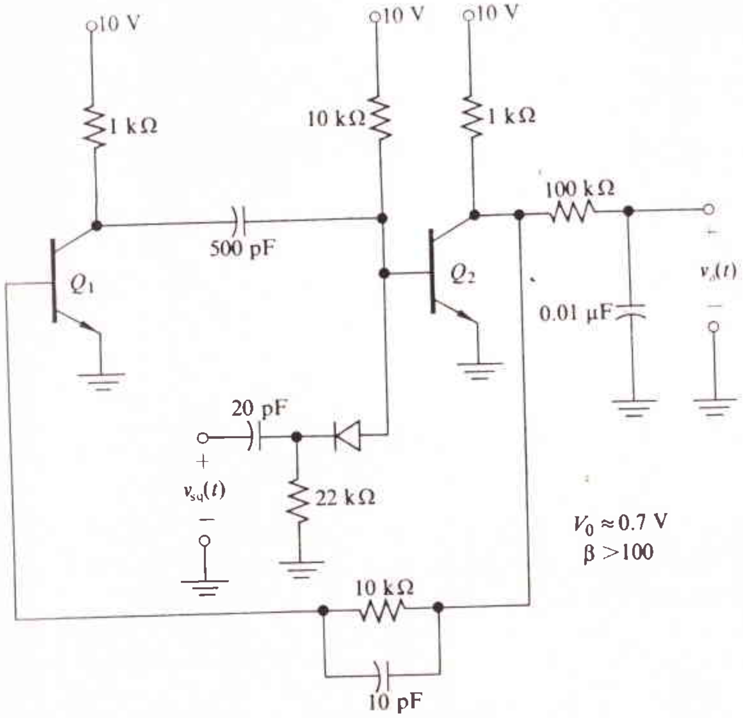


Figure 12.P–8

b) If $\omega_i(t) = (10^5 + 10^4 \cos 10^3 t)$ rad/sec, determine an expression for $v_o(t)$. What is the maximum carrier ripple which appears at $v_o(t)$?

12.20 Repeat Problem 12.19 for the case where the trigger circuit at the base of Q_2 in Fig. 12.P–8 is replaced by the trigger circuit shown in Fig. 12.P–9. What is the advantage of this more complicated circuit?

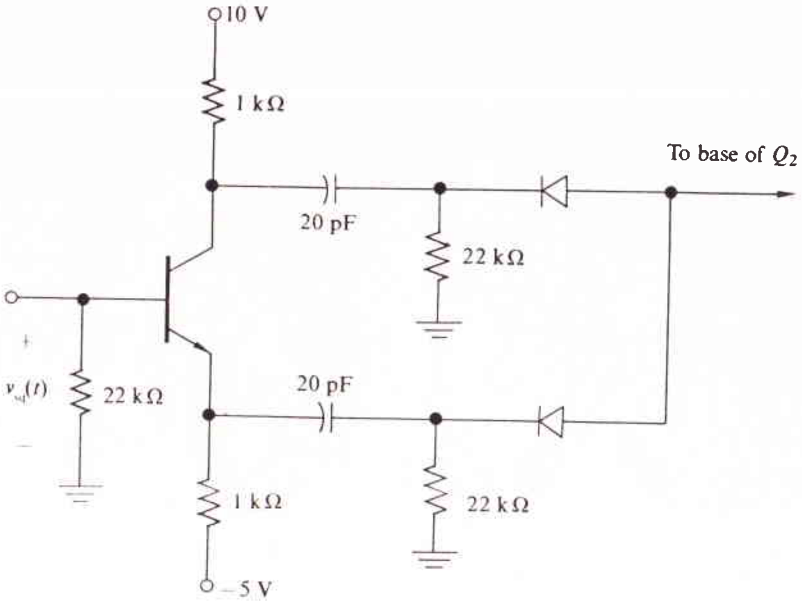


Figure 12.P–9

APPENDIX

MODIFIED BESSEL FUNCTIONS

The modified Bessel function of order n and argument x may be defined as the integral function given by

$$I_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{x \cos\theta} \cos n\theta \, d\theta. \tag{A–1}$$

A functional relationship between $I_n(x)$ and x may be achieved by integrating Eq. (A–1) numerically (on a computer or by hand) for various values of x. Table A–1 presents values of $I_n(x)$ for $n = 0, 1, 2, 3$ and values of x between 0 and 10. Table A–2 presents values of $I_n(x)$ normalized to e^x for $n = 0, 1, 2, 3, 4, 5$ and values of x between 0 and 20. In addition, Table A–3 presents values of $2I_n(x)/I_0(x)$ for the same values of n and x. It is quite clear that $I_n(x)$ is a monotonically increasing function of x.

For small values of x we may expand $e^{x\cos\theta}$ in a MacLauren expansion and keep only the first few terms in the series. With this approximation inserted in Eq. (A–1) we obtain

$$I_1(x) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + x\cos\theta)\cos\theta \, d\theta = \frac{x}{2}$$

and

$$I_2(x) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(1 + x\cos\theta + \frac{x^2\cos^2\theta}{2}\right) \cos 2\theta \, d\theta = \frac{x^2}{8}.$$

In a similar fashion we can show that, as $x \to 0$,

$$I_n(x) \to \frac{(x/2)^n}{n!}. \tag{A–2}$$

By keeping all the terms in the expansion for $e^{x\cos\theta}$ and performing a term-by-term integration, we can reduce Eq. (A–1) to an equivalent MacLauren series in x for $I_n(x)$. This series expansion is found in all the literature on modified Bessel functions.

For large values of x, $e^{x\cos\theta}$ has a very sharp dominant peak in the vicinity of $\theta = 0$ and may therefore be approximated by

$$e^{x\cos\theta} \approx e^{x(1-\theta^2/2)};$$

Table A–1 Tabulation of $I_n(x)$ vs. x for $n = 0, 1, 2, 3$

x	$I_0(x)$	$I_1(x)$	$I_2(x)$	$I_3(x)$
0.0	1.0000	0.00000	0.00000	0.00000
0.5	1.0635	0.25789	0.03191	0.00265
1.0	1.2661	0.56516	0.13575	0.02217
1.5	1.6467	0.98167	0.33783	0.08077
2.0	2.2796	1.59060	0.68895	0.21274
2.5	3.2898	2.51670	1.27650	0.47437
3.0	4.8808	3.9534	2.2452	0.95975
3.5	7.3782	6.2058	3.8320	1.82640
4.0	11.3020	9.7595	6.4222	3.33730
4.5	17.4810	15.3890	10.6420	5.93010
5.0	27.2400	24.3360	17.5060	10.33100
5.5	42.695	38.588	28.663	17.743
6.0	67.234	61.342	46.787	30.150
6.5	106.290	97.735	76.220	50.830
7.0	168.590	156.040	124.010	85.175
7.5	268.160	249.580	201.610	142.060
8.0	427.56	399.87	327.59	236.07
8.5	683.16	641.62	532.19	391.17
9.0	1093.6	1030.9	864.49	646.69
9.5	1753.5	1658.4	1404.30	1067.20
10.0	2815.7	2671.0	2281.50	1758.40

with this approximation, Eq. (A–1) reduces to

$$I_n(x) \approx \frac{e^x}{2\pi} \int_{-\pi}^{\pi} e^{-x\theta^2/2} \cos n\theta \, d\theta \approx \frac{e^x}{2\pi} \int_{-\infty}^{\infty} e^{-x\theta^2/2} \cos n\theta \, d\theta. \tag{A–3}$$

Equation (A–3) is directly integrable and yields

$$I_n(x) \approx \frac{e^x}{\sqrt{2\pi x}} e^{-n^2/2x} \tag{A–4}$$

$$\approx \frac{e^x}{\sqrt{2\pi x}} \left(1 - \frac{n^2}{2x}\right).$$

For $x \geq 10$, Eq. (A–4) is accurate within 1% for $I_0(x)$.

A–2 Tabulation of $I_n(x)/e^x$ vs. x for $n = 1, 2, 3, 4, 5$

x	$I_0(x)/e^x$	$I_1(x)/e^x$	$I_2(x)/e^x$	$I_3(x)/e^x$	$I_4(x)/e^x$	$I_5(x)/e^x$
0.0	1.0	0.0	0.0	0.0	0.0	0.0
0.5	0.64503	0.15642	0.01935	0.00160	0.00010	0.00005
1.0	0.46576	0.20791	0.04994	0.00816	0.00101	0.00010
1.5	0.36743	0.21904	0.07538	0.01802	0.00329	0.00048
2.0	0.30851	0.21527	0.09324	0.02879	0.00687	0.00133
2.5	0.27005	0.20658	0.10478	0.03894	0.01133	0.00270
3.0	0.24300	0.19683	0.11178	0.04778	0.01622	0.00454
3.5	0.22280	0.18740	0.11572	0.05515	0.02117	0.00676
4.0	0.20700	0.17875	0.11763	0.06112	0.02594	0.00924
4.5	0.19420	0.17096	0.11822	0.06588	0.03038	0.01187
5.0	0.18354	0.16397	0.11795	0.06961	0.03442	0.01454
5.5	0.17448	0.15770	0.11714	0.07251	0.03804	0.01719
6.0	0.16666	0.15205	0.11597	0.07474	0.04124	0.01975
6.5	0.15980	0.14694	0.11459	0.07642	0.04405	0.02220
7.0	0.15374	0.14229	0.11308	0.07767	0.04651	0.02452
7.5	0.14832	0.13804	0.11150	0.07857	0.04865	0.02668
8.0	0.14343	0.13414	0.10990	0.07919	0.05050	0.02869
8.5	0.13900	0.13055	0.10828	0.07959	0.05210	0.03056
9.0	0.13496	0.12722	0.10669	0.07981	0.05348	0.03227
9.5	0.13125	0.12414	0.10512	0.07988	0.05467	0.03384
0.0	0.12783	0.12126	0.10358	0.07983	0.05568	0.03528
0.5	0.12467	0.11858	0.10208	0.07969	0.05655	0.03660
.0	0.12173	0.11606	0.10063	0.07947	0.05728	0.03780
.5	0.11899	0.11369	0.09922	0.07918	0.05790	0.03890
2.0	0.11643	0.11146	0.09785	0.07885	0.05842	0.03990
2.5	0.11402	0.10936	0.09652	0.07847	0.05886	0.04081
3.0	0.11176	0.10737	0.09524	0.07807	0.05921	0.04163
3.5	0.10963	0.10549	0.09400	0.07764	0.05950	0.04238
4.0	0.10761	0.10370	0.09280	0.07718	0.05972	0.04306
4.5	0.10571	0.10200	0.09164	0.07672	0.05990	0.04367
5.0	0.10390	0.10037	0.09051	0.07624	0.06002	0.04422
5.5	0.10218	0.09883	0.08943	0.07575	0.06011	0.04473
6.0	0.10054	0.09735	0.08838	0.07526	0.06015	0.04518
6.5	0.09898	0.09594	0.08736	0.07476	0.06017	0.04559
7.0	0.09749	0.09458	0.08637	0.07426	0.06016	0.04595
7.5	0.09607	0.09328	0.08541	0.07376	0.06012	0.04628
8.0	0.09470	0.09204	0.08448	0.07326	0.06006	0.04657
8.5	0.09340	0.09084	0.08358	0.07277	0.05998	0.04683
9.0	0.09214	0.08969	0.08270	0.07227	0.05988	0.04706
.0	0.09094	0.08858	0.08185	0.07179	0.05977	0.04727
.0	0.08978	0.08750	0.08103	0.07130	0.05964	0.04744

Table A–3 Tabulation of $2I_n(x)/I_0(x)$ vs. x for $n = 1, 2, 3, 4, 5$

x	$2I_1(x)/I_0(x)$	$2I_2(x)/I_0(x)$	$2I_3(x)/I_0(x)$	$2I_4(x)/I_0(x)$	$2I_5(x)/I_0(x)$
0.0	0.0	0.0	0.0	0.0	0.0
0.5	0.4850	0.0600	0.0050	0.0003	0.0000
1.0	0.8928	0.2144	0.0350	0.0043	0.0004
1.5	1.1923	0.4103	0.0981	0.0179	0.0026
2.0	1.3955	0.6045	0.1866	0.0445	0.0086
2.5	1.5300	0.7760	0.2884	0.0839	0.0200
3.0	1.6200	0.9200	0.3933	0.1335	0.0374
3.5	1.6822	1.0387	0.4951	0.1900	0.0607
4.0	1.7270	1.1365	0.5906	0.2506	0.0893
4.5	1.7607	1.2175	0.6785	0.3129	0.1222
5.0	1. 768	1.2853	0.7585	0.3751	0.1584
5.5	1.8076	1.3427	0.8311	0.4360	0.1970
6.0	1.8247	1.3918	0.8969	0.4949	0.2370
6.5	1.8390	1.4342	0.9564	0.5513	0.2779
7.0	1.8511	1.4711	1.0104	0.6050	0.3189
7.5	1.8615	1.5036	1.0595	0.6560	0.3598
8.0	1.8705	1.5324	1.1043	0.7042	0.4001
8.5	1.8784	1.5580	1.1452	0.7497	0.4396
9.0	1.8854	1.5810	1.1827	0.7926	0.4782
9.5	1.8916	1.6018	1.2172	0.8330	0.5157
10.0	1.8972	1.6206	1.2490	0.8712	0.5520
10.5	1.9022	1.6377	1.2784	0.9072	0.5872
11.0	1.9068	1.6533	1.3056	0.9412	0.6211
11.5	1.9110	1.6677	1.3309	0.9733	0.6538
12.0	1.9148	1.6809	1.3545	1.0036	0.6854
12.5	1.9183	1.6931	1.3765	1.0324	0.7157
13.0	1.9215	1.7044	1.3970	1.0596	0.7450
13.5	1.9244	1.7149	1.4163	1.0854	0.7731
14.0	1.9272	1.7247	1.4344	1.1099	0.8002
14.5	1.9298	1.7338	1.4515	1.1332	0.8262
15.0	1.9321	1.7424	1.4675	1.1554	0.8513
15.5	1.9344	1.7504	1.4827	1.1765	0.8754
16.0	1.9365	1.7579	1.4970	1.1966	0.8987
16.5	1.9384	1.7650	1.5105	1.2158	0.9211
17.0	1.9403	1.7717	1.5234	1.2341	0.9426
17.5	1.9420	1.7781	1.5356	1.2516	0.9634
18.0	1.9436	1.7840	1.5472	1.2683	0.9835
18.5	1.9452	1.7897	1.5582	1.2843	1.0028
19.0	1.9466	1.7951	1.5687	1.2997	1.0215
19.5	1.9480	1.8002	1.5788	1.3144	1.0395
20.0	1.9493	1.8051	1.5883	1.3286	1.0569

To obtain derivatives of $I_n(x)$, we differentiate both sides of Eq. (A–1) with respect x to obtain

$$\frac{dI_n(x)}{dx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos\theta\, e^{x\cos\theta} \cos n\theta\, d\theta$$

$$= \frac{1}{2}\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{x\cos\theta}\cos(n+1)\theta\, d\theta + \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{x\cos\theta}\cos(n-1)\theta\, d\theta\right] \qquad \text{(A–5)}$$

$$= \frac{1}{2}\left[I_{n+1}(x) + I_{n-1}(x)\right].$$

quation (A–5) is valid for $n \geq 1$. For $n = 0$ we obtain in an identical fashion

$$\frac{dI_0(x)}{dx} = I_1(x). \qquad \text{(A–6)}$$

rom Eq. (A–1) it is apparent that

$$e^{x\cos\theta} = I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)\cos n\theta. \qquad \text{(A–7)}$$

we wish to evaluate the Fourier series of $e^{x\sin\theta}$, we may define $\phi = \theta - \pi/2$ and bserve that

$$e^{x\sin\theta} = e^{x\cos\phi} = I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)\cos n\phi$$

$$= I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)\cos\left(n\theta - \frac{n\pi}{2}\right)$$

$$= I_0(x) + 2\sum_{n=1}^{\infty} (-1)^n I_{2n}(x)\cos 2n\theta$$

$$+ \sum_{n=1}^{\infty} (-1)^{n+1} I_{2n-1}(x)\sin(2n-1)\theta. \qquad \text{(A–8)}$$

a similar fashion we can show that

$$e^{-x\cos\theta} = I_0(x) + 2\sum_{n=1}^{\infty} (-1)^n I_n(x)\cos n\theta \qquad \text{(A–9)}$$

defining $\phi = \theta - \pi$.

We may obtain a recursion formula for the modified Bessel functions by differentiating both sides of Eq. (A–7) with respect to θ:

$$2\sum_{n=1}^{\infty} nI_n(x)\sin n\theta = x\sin\theta e^{x\cos\theta}$$

$$= x\sin\theta\left[I_0(x) + 2\sum_{n=1}^{\infty} I_n(x)\cos n\theta\right]$$

$$= x\sin\theta I_0(x) + x\sum_{n=1}^{\infty} I_n(x)\sin(n+1)\theta$$

$$- x\sum_{n=1}^{\infty} I_n(x)\sin(n-1)\theta. \qquad \text{(A–10)}$$

By equating the coefficients of the corresponding ($\sin n\theta$)-terms in the first and last terms of Eq. (A–10), we obtain the desired result,

$$2nI_n(x) = x[I_{n-1}(x) - I_{n+1}(x)]. \qquad \text{(A–11)}$$

Additional relationships involving Bessel functions can be found in the following references:

1. F. E. Relton, *Applied Bessel Functions*, Blackie and Son, Limited (1946), Chapter 3.
2. G. N. Watson, *Theory of Bessel Functions*, Cambridge University Press, Cambridge (1922).
3. E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*, Cambridge University Press, Cambridge (1927), Section 17.7.

Additional tabulations and relationships of modified Bessel functions may be found in the following references:

1. M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematical Series No. 55, Government Printing Office, Washington D.C. (1964). Reissued by Dover, New York (1965). Section 9.6 and Tables 9.8 through 9.11 deal with modified Bessel functions. Tabulated values as high as $I_{100}(100)$ are given.
2. E. Jahnke and F. Emde, *Tables of Functions*, fourth edition, Dover, New York (1945). This standard reference work tabulates $I_0(x)$ and $I_1(x)$ for values of x up to 9.99. It also contains other functions up to $I_{11}(6)$, $I_0(x) = J_0(jx)$, $I_1(x) = -jJ_1(jx)$, etc.
3. M. Onoe, *Modified Quotients of Bessel Functions*, Columbia University Press, New York (1958). This work tabulates $xI_0(x)/I_1(x)$ for values of x up to 20. The reciprocal of this function is within a factor of two of $2I_1(x)/xI_0(x)$, which is the ratio of G_m to g_m.

ANSWERS TO SELECTED PROBLEMS

CHAPTER 1

1.1. $P_1 = 4.15\ \text{mW}, \qquad P_2 = 1.65\ \text{mW}, \qquad P_3 = 0.55\ \text{mW}$

1.6. $v_o(t) \approx -0.28 \sin 10^7 t + 7.3 \cos 2 \times 10^7 t + 0.33 \sin 3 \times 10^7 t$

1.7. $v_o \approx -5.9\ \text{V}, \qquad V_{EB_1} \approx 815\ \text{mV}$

1.8. $v_o(t) \approx 118\ \text{mV} \cos 10^6 t$

CHAPTER 2

2.1. $n = 19.1; \qquad f_{\text{High}} = 25.8\ \text{kHz}, f_{\text{Low}} = 266\ \text{Hz}; \qquad v_2/v_1 = 1/20.9; \eta = 0.914$

2.7. Bandwidth is 242 kHz

2.8. $P_{\text{load}} = 4.6\ \text{W}$

2.10. Loaded Q is 16.2

CHAPTER 3

3.1.
$$Y_{11L}(p) \approx \frac{1/2L}{p + (1/2GL)}$$

This may be represented physically as the series combination of a conductance G and an inductor $2L$.

3.2. $i_2(t) = 100\ \text{mA}\left[1 + \frac{1}{\sqrt{2}} \cos\left(2.5 \times 10^5 t - \frac{\pi}{4}\right)\right]$

3.3. a) $10^9\ \text{V}e^{-10^6 t/3} \cos 10^7 t\, u(t)$

b) $100\ \text{V}e^{-10^6 t/3} \sin 10^7 t\, u(t)$

c) $1500\ \text{V}[1 - e^{-10^6 t/3}] \cos 10^7 t\, u(t)$

The voltages appear large since unit drives were assumed; i.e., an impulse of one coulomb strength and a step with a one-ampere level.

3.6. $V_{o22} = 200\, I_2/\sqrt{5}$ (cf. Ex. 2.4-3); $I_2 = 2.67\ \text{mA}$

3.7. $v_o(t) = \sqrt{2}\ \text{V} \cos\left(2.5 \times 10^5 t - \frac{\pi}{4}\right) \cos 10^7 t$

CHAPTER 4

4.2. Fundamental is 0.39 mA; third harmonic is 0.145 mA.

4.6. $G_m = 400\ \mu\text{A/V}$ when $V_{dc} = 3\text{V}$

4.7. $I_2 = 44\ \mu\text{A}$

4.8. $I_1 = 8.45\ \text{mA}, I_2 = 7.8\ \text{mA}$

4.11. $V_{\text{tank}} = 9.3\ \text{V}$

4.12. $P_{\text{battery}} = 27\ \text{mW}$

HAPTER 5

2. $P_{lk} = 12.5$ mW
5.3. $V_C \approx 540$ mV
5. $V_c \approx 690$ mV
5.6. $C \geq 63\ \mu$F; conduction angle $\approx 120°$.

HAPTER 6

1. a) $A_{\min} = \sqrt[3]{8}$, $C = 1370$ pF; b) $|V_3(\omega)| = |V_i(\omega)| = 1$ V
2. b) $|V_3(\omega)| = 0.535$ V with no change in frequency.
3. $\omega_0 = 10^7$ rad/sec, $|V_x(\omega)| = 4$ V, $|I_y(\omega)| = 200\ \mu A$
5. All three poles originate at $1/2RC$ for $\mu = 0$ and move outward along radial paths separated by 120° as μ is increased. One path lies on the negative real axis. For $\mu = 4\sqrt[3]{8}$, two of the poles cross the imaginary axis at $\omega_0 = \sqrt{3}/2RC$.
7. $\omega_0 = 10^4$ rad/sec, $|V_1(\omega_0)| = 1.66$ V. (Note that the attenuation through each cathode follower is 0.815.)
9. $A_{\min} = 9$, $\omega_0 = 1/\sqrt{5}RC$
6.10. $A_{\min} = +93$, $S_F = 20$
11. $r = 25\ \Omega(1 - 2/A) = 24.75\ \Omega$, -0.268 rad/sec
12. 2.36 mA, 1.34 mA
13. $v_o(t) = 12.5$ V $\cos 4.5 \times 10^6 t$; $THD = 0.37\%$

HAPTER 7

1. $i_{\text{difference}} = 450$ nA, $i_{\text{signal}} = 905$ nA, $i_{\text{oscillator}} = 1.62$ mA
3. $i_{\text{difference}} = 66\ \mu$A, $i_{\text{oscillator}} = 1.72$ mA
5. $G_c \approx 420$ mhos
9. $v_o(t) = g(t)[860] \cos(\omega_1 - \omega_2)t$

HAPTER 8

1. b) 5 V at ω_0; 1.5 V at $(\omega_0 + \omega_m)$ and at $(\omega_0 - \omega_m)$
2. b) 125 mW; 22.5 mW
3. b) $5\text{ V} + 2.12\text{ V} \cos\left(\omega_m t - \frac{\pi}{4}\right) \cos \omega_0 t$
4. a) $\hat{g}(t) = (5\text{ V}) \sin \omega_1 t + (10\text{ V}) \sin \omega_2 t$
7. $-1.06\text{ V}\left[1 + 0.707 \cos\left(10^4 t - \frac{\pi}{4}\right)\right] \cos(3 \times 10^7 t)$
8. 0.495
9. $v_o(t) = (750\text{ mV})[1 + 0.5 f(t)] \cos 10^7 t$
17. $v_o = (10\text{ V}) - 2\text{ V}\left[1 + 0.354 \cos\left(10^5 t - \frac{\pi}{4}\right)\right] \cos 10^7 t$

HAPTER 9

2. 48°C with no ac, 29°C with ac
5. $\eta = 0.30$, $\eta = 0.60$, $\eta = 0.81$
7. $P_{\text{transistor}} \approx 7$ W
9.10. $\eta \approx 0.77$

CHAPTER 10

10.1. $v_o(t) = 7.5\text{ V} - (45\text{ mV}) \cos\left(10^3 t - \frac{\pi}{4}\right)$

10.3. $v_o(t) = \frac{A^2}{2}[1 + mf(t)]^2 + AB[1 + mf(t)] + \frac{B^2}{2} \quad * h_L(t)$

10.8. $(40\text{ mV}) \cos 10^4 t$

10.9. $(3.18\text{ V})\left[1 + 0.707\text{m} \cos\left(10^4 t - \frac{\pi}{4}\right)\right]$

10.10. $v_i(t) = (100\text{ V}) \cos 10^7 t$; $v_o(t) \approx 100$ V, $\phi = 14.4°$; $FR = 0.0157$
10.11. For large Q_T, $THD \approx 0.47/Q_T$
10.12. a) $v_o = 62.5$ V, $v_i = (62.5\text{ V}) \cos 10^7 t$, $Q_T = 125$, $FR = 0.025$, $\varphi = 18.2°$, $THD = 0.0075$
b) $v_o' = 125$ V, $v_o \approx 0$, $v_i = (125\text{ V}) \cos 10^7 t$, $Q_T = 250$, $FR = 0{\cdot}0125$,
$\varphi = 13.2°$, $THD = 0.0038$

10.14. $v_o(t) = (44.2\text{ V}) \cos\left(2 \times 10^4 t - \frac{\pi}{4}\right)$

$v_o(t) = (125\text{ V}) + (44.2\text{ V}) \cos\left(2 \times 10^4 t - \frac{\pi}{4}\right)$

$v_i(t) = (125\text{ V})\left[1 + 0.353 \cos\left(2 \times 10^4 - \frac{\pi}{4}\right)\right] \cos 10^7 t$

10.15. $R = 12.5\text{ k}\Omega$

10.17. $v_1(t) = -(3.75\text{ V})\left[1 + 0.471 \cos\left(10^4 t - \frac{\pi}{4}\right)\right] \cos 10^7 t$

$v_2(t) = (3.75\text{ V}) + (1.77\text{ V}) \cos\left(10^4 t - \frac{\pi}{4}\right)$

$v_3(t) = (0.885\text{ V}) \cos\left(10^4 t - \frac{\pi}{4}\right)$

CHAPTER 11

11.2. $BW = 1.6 \times 10^6$ rad/sec
11.3. $\Delta f = 625$ kHz

11.4. $v_o(t) = (11\text{ V}) - (5\text{ V}) \cos\left(2 \times 10^8 t + 2 \times 10^5 \int^t \cos 10^4 \theta \, d\theta\right)$

11.5. -57.7 dB

11.7. $v_o = (100\text{ mV}) \sin\left(10^7 t + 5 \times 10^6 \int^t f(\theta)\, d\theta\right)$

11.10. $\omega_0 = 10^7$ rad/sec, $\Delta\omega \approx 10^4$ rad/sec
11.11. $\omega_i(t) = (2.5 \times 10^7\text{ rad/sec})\,[1 + 0.1 f(t)]$

$v_i(t) = 2\text{ V}\cos\left(10^7 t + 1.25 \times 10^{-3}\int^t \cos 10^3\theta\, d\theta\right)$

$f_i(t) = 10^5\text{ Hz}[1 + \frac{1}{2}f(t)]$

TER 12

$V_1 \approx 0.8\text{ mV}, m_0 \approx 0.02$

$BW > 5.3\,\Delta\omega, D < 0.027$

12.4. $BW > 14\Delta\omega, D < 0.0028$

12.7. 0.0075, 0.075, no.

$v_o(t) = (5/\sqrt{2})\text{ V}\cos\left(10^4 t - \frac{\pi}{4}\right)$

a) $1570\ \Omega$, d) $v_o(t) = (250\text{ mV})\cos 10^3 t$

$v_o(t) = 0.55\text{ V}\cos\omega_m t$

$v_o(t) = \left(\frac{5\text{ V}}{\sqrt{2}}\right)(1 + 0.2\cos 10^5\, t)$

b) $v_o(t) = (0.58\text{ V}) + (0.041\text{ V})\cos\left(10^3 t - \frac{\pi}{4}\right)$

INDEX

INDEX

Abramowitz, M., 641
AFC, 590
AGC, 308, 331–336, 486, 530
Alignability, 324–327
AM generators: balanced, 382–384
  differential amplifier, 374–375
  diode bridge, 376–379, 382–384
  FET, 380–381
  integrated circuit, 395–396
  square law, 384–387
  tuned circuit, 387–392
  See also Diode bridge, FET, Chopper modulation, Multiplier, *and/or* Modulation
AM methods: analog, 354–355
  chopper, 355–358
  direct, 361–362
  high level, 447–450
  nonlinear, 359–360
AM on FM signals, 589–590
AM receiver, superheterodyne, 478
AM signals: in capacitively coupled circuits, 168–169, 195–199
  in narrowband filters, 72, 76, 88–89
  normal, 347
  single sideband, 350–352, 457–461, 464–466
  suppressed carrier, 349, 357
Amplification, AM, 405–406
Amplifiers: cascode, 324–326
  class A, 401–405
  class AB, 401
  class B, broadband, 401, 421–425
  class B "linear," 92, 405–408
  class C "linear," 408–410
  class C–RF, 401, 410–415
  class D, broadband, 426–432
  class D, narrowband, 401, 415–421
  differential pair, 114–119, 324–326
  feedback, 206, 475–478
  FET, 98–104, 131–136
  gain controlled, 214, 335–336
  maximum signal without distortion, 4, 100, 110
  narrowband-tuned, 6–8, 314–331, 405–421
  power, 401–455
  small signal range, 4–5, 110, 118
  transistor, 1–8, 104–113
  tuned, 6–8, 314–331, 405–421
Amplitude control, 212–215, 265–266, 308, 331–336
Amplitude limiting, 7–8, 110, 188, 571–577
  nonlinear, 207
Amplitude modulation; *see specific headings under* AM
Amplitude sensitivity, function, 230–231, 264–265
Amplitude stabilization: FM, 530–532
  oscillators, 212–215, 262–266
Analog computer, 528
Angelo, E. J., Jr., 13
Armstrong, E. H., 560
Astable multivibrator, 553–556
Attenuator, nonlinear, 213, 265–266
Auto transformer, 49–51
Automatic frequency control (AFC), 590
Automatic gain control; *see* AGC
AVC; *see* AGC
Average envelope detection, 461–466
  in FM detectors, 582–584, 587
Average envelope detector, 468–478
  dual transistor, 473–475
  operational amplifier, 475–478
Avins, J., 613

ck diode, 505–506
ghdady, E. J., 523, 526
ndwidth: FM, 513–515, 565
  parallel RLC circuit, 28
  peak envelope detector, 485, 494–495
  pi network, 441–443
  power amplifier coupling network, 433–437
  transformer, 21
rkhausen criterion, 11, 208
ennett, W. R., 628
essel functions: modified
    in amplifiers, 108–111
    approximations, 113, 636
    in mixers, 293–300
    tabulations, 109, 637–639
    theory, 636–641
    uses, 3–9, 108–111, 293–300
  ordinary, 427, 511–514
    plot of, 512
    zeros of, 514
ias shifts, 4–5
  in transistor circuits, 149, 169–181
iasing: clamp circuit, 131–134
  constant current, 1–2, 162–169
  resistive, 169–181
  triode grid, 432–433
ilotti, A., 618
inary AM, 410
oltzmann's constant, 2

admium sulphide, 560
alandrino, L., 628
alibration: FM generator, 514
  high-Q filter circuits, 88–89
apacitive coupling, 162–171
  to nonlinear devices, 149–180
apacitive step-down circuits, 38–48
arrier, pilot, 353
arrier cancellation, 355, 376–383
arrier filtering, 459–461
ascode amplifier, 324–327, 345–346
assara, F., 628
hopper modulation, 355–358
hudobiak, W. J., 421
lamping circuits, 131–134, 153–180, 188
  current-driven, 153–155
  inductively biased, 154–156
  peak FET, 131–134, 301–302, 304–305

Clarke, K. K., 401, 588, 628
Clarke-Hess FM detector, 587–592, 632
Collector saturation in oscillators, 234–236
Colpitts oscillator, 223, 226, 231, 253
Complementary pair transistors, 423
Conduction angle, 93–103, 107
  peak envelope detector, 482
Conversion transconductance; *see* Transconductance
Converters, transistor, 311–314
Convolution, 74, 581
Coupling networks, 16–64, 433–447
Cross modulation, 123, 337–338
Crystal equivalent circuit, 243–247
Crystal filter, 459
Crystal FM detector, 564, 602, 632
Crystal oscillator, 243–255
Current-driven junction, 249
Current inverter, 547–549
Current pulse, peak envelope detector, 480–482
Current-switching circuit, 417–420, 431–432

Dammers, B. G., 260
Delay line, 603, 634
  helical, 609
Delay network: ideal, 604
  RLC, 610
Demodulation: amplitude, 457–508
  frequency, 571–635
Derivative, definition, 602
Deviation ratio, 511
Diagonal clipping; *see* Failure-to-follow distortion
Differential amplifier; *see* Differential pair
Differential pair: amplifiers, 114–119
  amplitude modulator, 365–375
  distortion in, 114–119
  mixers, 309–310
  multiplier, 365–375
  oscillator, 236–240, 251–252, 288, 291
  RF amplifiers, 324–326
Differentiating network: direct, 586–587
  frequency domain, 522, 586, 593
  time delay, 586, 602–603
Diode bridge, 376–379
  modulator, 376–383

Diode clamping circuit, 153–162
Diode detector, 468–503
Diode shaping network, 551–553
Diodes: back, 505–506
  germanium, 471, 586
  hot carrier, 553
  silicon, 2, 471, 496, 499
  variable capacitance, 541–542
  Zener, 557
Distortion: in AM, 361–362
  failure-to-follow; *see* Failure-to-follow distortion
  in FM detectors, 595–623
  phase, 390
  total harmonic; *see* Total harmonic distortion
Double-tuned circuits, 328–331
Double-tuned transformers, 433–439
Drain resistance (in FET's), 134–136

Edson, W. A., 244
Efficiency power amplifiers, 402ff
Efficiency transformer, 19, 402
Emde, F., 641
Enloe, L. H., 628
Envelope detector, 468; *see also* Average envelope detector, Peak envelope detector
Envelope detector loading, 489
  in FM demodulators, 600, 610
Equivalent circuits: low pass, 65–81
  peak envelope detector, 485
  transformer, 51, 56, 62–64
Equivalent linear resistance, 181–194
  envelope detector, 481
Exponential characteristic, 104–113

Failure-to-follow distortion, 361, 467, 389–390, 577
  in amplitude modulators, 389–390
  in broadband peak detector, 500
  in peak narrowband detector, 490–494
Federal Communications Commission, 562
Feedback: amplifier, 206, 475–478
  detector, 475–478
  FM detector with, 623–628
  FM generator stabilization, 562–565
FET: AGC circuit, 266
  amplifiers, 98–104, 131–136
  clamp-biased circuits, 131–134
  distortion in, 98–104
  drain resistance, 134–136
  mixer, 295–297, 301–305
  multiplier, 363–365
  oscillator, 241–243, 253–255
  switch, 380–381
Filaments, tungsten, 414
Filter: crystal, 459
  mechanical, 315–316
  pulse-forming, 620
FLL, 623–625
FM demodulation, 571ff
FM demodulator: balanced slope, 596–600
  balanced time delay, 608–613
  balancing in, 589–591
  Clarke-Hess, 587–592, 632
  crystal, 564, 602, 632
  direct differentiation, 586–588
  FMFB, 623–626
  Foster-Seeley, 609–613
  frequency locked loop (FLL), 623–626
  ideal, 578
  laser, 608, 633
  microwave, 608
  phase locked loop (PLL), 623–628
  pulse count, 634, 635, 618–623
  ratio detector, 613–615
  series tuned slope, 600–602
  simplified time delay, 615–618
  slope, 593–602
  time delay, 602–623
FM detectors; *see* FM demodulators
FM differential equation, 527–528
  quasi-static, 532–533
FM generators: analog, 528–529
  Armstrong, 527, 559–562
  crystal-controlled, 560
  differential pair, 537
  FET, 569–570
  laser, 560, 562
  quasi-static, 534–542
  single transistor, 540
  square wave, 553–559
  time delay, 562
  triangular wave, 568, 549
  variable capacitance, 541–542
FM noise threshold, 624
FM signals: impulse train, 579, 584
  nonsinusoidal, 520–521

UNIVERSIDAD DE LA REPUBLICA
FACULTAD DE INGENIERIA
DEPARTAMENTO DE DOCUMENTACION Y BIBLIOTECA
MONTEVIDEO - URUGUAY

square, 577, 520–521
triangular, 520–521, 542–545, 632
sawtooth modulated, 510
sidebands of, 513
M spectrum, 511–515
MFB, 623–625
ster, D. F., 610
ster-Seeley detector, 609–613
ourier-like" FM expansion, 516
urier series: half-sine wave train, 92
impulse train, 80
narrow-pulse train, 480
periodic signal, 79
rectangular pulse train, 408
sine-wave tips, 93–95, 144–148
square wave, 92
switching function, 355–356, 462
urier transform, 72
of AM signal, 72–73, 347–349
of Hilbert transformer, 351
of nonlinear FM, 517
of unit step, 71
equency demodulator with feedback (FMFB), 623–625
equency depression, 279, 282
equency deviation, definition, 510
equency division multiplexing (FDM), 73, 509
equency doubler, 561
equency locked loop, 623–625
equency modulation, 509–570; *for breakdown, see specific topics under* FM
equency stability, 210, 216–222
direct, 216–217
FM generators, 562–565
indirect, 217–222
equency tripler, 561
utiger, P., 628
ll-wave rectifier, 503

ain control, 265–266, 308, 331–336; *see also* AVC
rber, E. A., 243
ermanium diode, 471, 496, 499
ilbert, B., 365
ray, P. E., 13
roszkowski, J., 244, 279

Haantjej, J., 260
Hafner, E., 243
Half-power frequency, 27
Half-wave rectifier, 462, 468–478, 618, 620
  operational amplifier, 475–478
Hall effect multiplier, 362–363
Hayashi, C., 273
Hess, D. T., 527, 588, 625, 628
Hilbert transform, 351–353, 457
"Hold in range," 626, 628
Hot carrier diodes, 553
Hysteresis, 544–545

IF amplifiers, 314–331
Immovilli, G., 628
Impedance matching, 16–19, 326–327
Impulse response, narrowband filter, 70
Input impedance, variation with frequency:
  in single-tuned circuits, 321–325
  in double-tuned circuits, 330–331
Instantaneous frequency, definition, 510
Instantaneous phase, definition, 509
Intermodulation, 242

Jahnke, E., 641
Joyce, M. V., 401
Junction voltage, 2, 4

Kirchhoff's current law, 480
Klein, G., 551

Large signal transconductance; *see* Transconductance, large signal
Lienard method, 273
Light bulb, 212, 213, 261
Limiters, 6–8, 110, 188, 571–577
  differential pair, 573–574, 599, 629–631
  diode, 572
  dynamic, 575–577, 613–614
  tuned circuit, 629–630
Linearization: exponential characteristic, 119, 123–129
  mixers, 302–305
Lissajous pattern, 88–89, 362
Load line, 183–190; *see also* Operating path
Local oscillator, 294
  frequency modulated, 590

Loop gain, 206
Loudspeaker, 421
Low pass equivalents, 65–69
  IF strip, 69

MacLean, W. R., 260
Matching, impedance, 16–19
Maximum modulation frequency, 511
Meacham bridge oscillator, 212, 261–267
Measurements, high Q circuits, 88–89
Mechanical filters, 315–316
Midband range transformer, 17, 21
Miller capacitance, 52
Miller effect, 241
Millman, J., 546, 553, 609
Mixers: differential-pair, 309
  FET, 295–297, 310–311
  general problems, 294–295, 300–302, 305–306
  series resistance in, 302–305
  transistor, 8–10, 297–300, 307
Model transformer, 17, 51, 63–64
Modified Bessel functions: in amplifiers, 108–111
  approximations, 113, 636
  in mixers, 293–300
  tabulations, 109, 637–639
  theory, 636–641
  uses, 3–9, 108–111, 293–300
Modulation, amplitude, 347–399; *for breakdown, see specific topics under* AM
Modulation, frequency, 509–570; *for breakdown, see specific topics under* FM
Modulation index: AM, 8, 77
  FM, 511
Monostable multivibrator, 622, 634–635
Multiplication of frequency deviation, 516
Multiplier: FET, 363–365, 379–381, 384–387
  four quadrant, 368–373
  frequency, 8
  integrated circuit, 365–373
  single quadrant, 367
  square law, 354
  transconductance, 395
  two quadrant, 368–369

Multivibrator: astable, 553–556
  monostable, 622, 634–635

Narrowband amplifiers, 6–7, 65–80, 112–119, 314–331
Narrowband networks, 65–81
Narrowband power amplifiers, 405–420, 432–447
Negative resistance, 205, 270–272
Networks: capacitive step-down, 38–48
  coupling, 433–439
  low pass equivalent, 65–69
  narrowband, general, 65–81
  pi, 439–447
  "transformer-like," 51
  transient response, 21–23, 70–72
Neutralization, 321, 327
Noise, 509
Noise figure, 336–337
Noise immunity in FM receivers, 571, 590
Noise threshold in FM systems, 624–629

Onoe, M., 641
Operational amplifier, 262
  half-wave rectifier, 475–478
  low-pass filter, 478
Operating paths: class A amplifier, 407
  class B amplifier, 407
  class C amplifier, 410
  effect of harmonics on, 438
Oscillators: Colpitts, 223, 226, 231–232
  crystal, 243–255, 262
    alignment, 251
    tuning, 246
  differential pair, 236–240, 251–252, 288, 291
  Hartley, 223
  junction FET, 241–243
  large signal model: differential pair, 237ff
    single transistor, 225ff
  local, 10, 233, 305–307
  Meacham bridge, 212, 261–267
  Miller, 253, 255
  negative-resistance, 268–272, 292
  one-port, 268–272
  pentode, 289
  phase shift, 212, 283–284
  Pierce, 253–254

  series mode crystal, 248–254
  sine wave, 10–12, 205ff
  single transistor, 222–236, 286–287, 289–290
  variable frequency (VCO), 233–234, 624–628
  Wien bridge, 212, 261, 267–268
Otte, J., 260
Overtone, crystal, 245

Page, D. F., 421
Parallel resonant circuits, 25–38
Peak envelope detection, 467–468
  in FM detectors, 584–586
Peak envelope detector, 532
  broadband, 498–503
  ideal, 467
  narrowband, 478–497
Pentodes: characteristics, 120–123
  mixers, 300–301
  oscillators, 253–255
  transconductance, 121–123
Periodic drives: effect of narrowband filters, 78–85
  effect of nonlinear elements, 91–134
Phase deviation, definition, 509
Phase distortion, 390
Phase locked loop (PLL), 459–460, 623–628
Phase modulation, 509–510
Phase plane, 273–279
Phasitron, 527
Pi network, 439–447
Piecewise linear models, 91, 124–126, 151–153
Pierce oscillator, 253–254
PLL, 459–460, 623–628
Pole zero diagram: delay network, 611
  double-tuned transformer, 435
  oscillator networks, 207, 208, 210, 212
  quartz crystal, 245
Power in AM signals, 349
Power amplifiers: class A, 401–405
  class AB, 401
  class B, broadband, 401, 421–425
  class C, linear, 408–410
  class C, RF, 401, 410–415, 442–447
  class D, broadband, 415–421
  class D, narrowband, 426–432
  coupling networks, 442–447
  modulation of, 387–391, 447–451
  switched, broadband, 415–421
  switched, narrowband, 426–432
Power gain, 326–327
Power loss in transformers, 17–19
"Pull in range," 628
Pulse count FM detector, 618–623, 634–635

Q: to limit output distortion, 82, 85
  parallel circuits, 27–28, 31
Quartz crystal, 210, 217
  model, 244
  oscillators, 243–255
  parameters, 245
Quasi-static approximation, 521–526, 596, 598
Quasi-static frequency modulation, 532–542

Ratio detector, 613–615
Raysistor, 266
Reactance tube, 542
Receiver design, 293–295
Relton, F. E., 641
Resonant frequency, 26–27, 33
Rice, S. O., 628
Ripple, 467
  in broadband envelope detector, 500
  in FM detectors, 592
  in peak envelope detector, 483
Root locus, 207–212, 434–437
Rowe, H. E., 426

Saturating collector, AM modulator, 388–392
Saturation current, transistor or diode, 2, 104, 171
Sawtooth generator, 428
Schmitt trigger, 543–547
Schwartz, M., 628
Seeley, S. W., 610, 613
Semiconductor resistor, 212
Series resistance: effect on clamping operation, 156–159
  effect on exponential characteristic, 123–130

  effect on mixer characteristic, 302–305
Signal generator, laboratory, 213
Signals: telephony, 505
  see also AM signals *and* FM signals
Silicon diode, 2, 471, 496, 499
Sine wave tips: with piecewise discontinuities, 92–98
  in power amplifiers, 414–415, 454–455
  with square law characteristic, 102–104
Single sideband (SSB): amplification, 405
  demodulation, 457–461, 464–466
  generation, 349–353, 368
  spectrum, 350
Small signal operation, 3, 101, 110, 118
Smith, W. L., 602
Smith Chart, 444–446
Source impedance, effect on clamping operation, 156–159
Spectrum: FM signal, 513
  SSB, 350
  video signal, 465
Spikes
Square law characteristics, 98–104, 131–134, 146–148
Square law detector, 500
Square law modulation, 359–360
Squegging, 255–261, 532
Stegun, I. A., 641
Stein, S., 628
Step response: narrowband networks, 70–75
  transformer, 21–23
Strauss, L., 551, 553
Superheterodyne, 10, 293–295
Superposition, 90, 149, 153
Suppressed carrier AM, 349, 357
Switch, voltage-controlled, 547–549
Switching function, 355–356, 462
Switching power amplifiers: broadband, 415–421
  narrowband, 426–432
Sykes, R. A., 243
Symmetric networks, 66
Synchronous detection, 457–461
  in FM detectors, 580–582
Synchronous detector, 615

Taub, H., 546, 553, 609
Taylor Series of RLC circuit magnitude, 594–595, 598
Thermal runaway, 400
Thermal time constant, 213
Thermistor, 212–213
Threshold extension FM receivers, 623–628
Threshold sensor, 543–544
Time-delay FM detector, 602–623
Time-varying controlled source, 534
Total harmonic distortion (THD), 82–85, 192–193
  in differential pair oscillator, 236, 240
  in transistor oscillator, 227–230
Transconductance: differential pair, 116–119
  effect of series resistance, 123–130
  FET, 100–101, 133–134
  pentodes, 121
  small signal, 100, 105, 116, 121
  transistor, 105–112, 128–129
Transconductance, conversion, 9, 297–314
  bipolar transistor, 9, 297–301
  differential-pair, 309–310
  FET, 295–297
  peak-clamped FET, 304–305
  pentode, 301–302, 304–305
Transconductance, large signal: differential pair, 119
  exponential characteristic, 7, 11, 110
    with series resistance, 129–130
  FET's, 100, 122, 134
  pentodes, 122–123
  resistively coupled transistor circuit, 177–179
  square law characteristic, 100, 134
Transconductance multiplier, 395
Transfer function, 65
Transformer: broadband, 16–24
  three-winding, 54–59
Transformer-like networks, 38–48
  summary, 51
Transient buildup: clamping circuit, 159–162
  transistor bias circuit, 165–169
Transient response: narrowband circuits, 71–76
Transistor: as amplifier, 104–113
  average detector, 473, 586–592
  biasing, 1–2, 162–180

- collector modulation of, 387–392
- limiters, 573
- mixers, 297–300
- simple oscillator, 222–227
- *see also individual topics such as* FET, Oscillators, Amplifiers, Biasing, *etc.*

Trapezoidal pattern, 88, 362
Triangular wave frequency modulation, 542–553
Triangular wave pulse train, harmonics in, 454–456
Triode class C amplifier, 411–415
Tuned circuit, nonlinear loading, 181–194
Turn on bias, 425
Turns ratio, transformer, 18–19

Unkauf, M., 628

Vacuum tube: AGC circuits, 332–334
- characteristics, 120–123, 334, 413
- mixers, 300–301
- modulated power amplifiers, 450–451
- oscillators, 253–255
- power amplifiers, 410–415, 432–439
- transconductance, 120–123

Van Suchtelen, H., 260
Varactor, 542
Varicap, 541–542
VCO, 541–542, 624–628
Vestigial sideband, 464–466
"Virtual" ground, 363, 476
Voltage-controlled oscillator (VCO), 541–542, 624–628

Watson, G. N., 641
Waveform function, 165
Waveshaping network, sinusoidal, 551–553
Whittaker, E. T., 641
Woodward, P. M., 515
Woodward's theorem, 515

Y parameters, 315–346

Zener diode, 557